Biochem. J. (1972) **128**, 427–444
Printed in Great Britain

The Complete Amino Acid Sequence of a Mouse κ Light Chain

By J. SVASTI and C. MILSTEIN
Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

(*Received* 10 *January* 1972)

The complete amino acid sequence of the κ-chain of the mouse myeloma protein MOPC 21 was established. The protein was reduced and alkylated with iodo[2-^{14}C]acetic acid, and 21 tryptic peptides were isolated, mainly by paper electrophoresis and paper chromatography. Three large tryptic peptides (of 35, 36 and 42 residues), which were difficult to isolate in this manner, were obtained pure and in excellent yields by a combination of Sephadex G-50 gel filtration in 1% (w/v) NH_4HCO_3 and chromatography on a DEAE-cellulose column in ammonium acetate buffer, pH 8.1. Peptides overlapping the tryptic peptides were isolated from a chymotryptic digest. The chain is 214 residues long. Microheterogeneity of two peptides was observed and is believed to be due to deamidation. It was not excluded that such deamidation could occur in serum from which the protein was isolated. The sequence is compared with the sequences of two other mouse κ-chains, and with the human κ-chain basic sequences.

Amino acid sequence studies have revealed that both the heavy and light chains of immunoglobulins consist of a variable *N*-terminal portion (V-region) of 100–120 residues and a constant *C*-terminal portion (C-region) comprising the remainder of the polypeptide chain (reviewed by Milstein & Pink, 1970). For proteins of a given class, type and allele, the *C*-terminal portion is identical, whereas until the work of Weigert *et al.* (1970), no two proteins have been found to have identical V-regions.

Three basic sequences can be clearly defined for the V-regions of human κ-chains (Milstein, 1967) and for human heavy chains (Wang *et al.*, 1970), but further subdivisions for the human κ-chains have been suggested (Milstein & Deverson, 1971; Schiechl & Hilschmann, 1971; Smith *et al.*, 1971). Subgroups have also been described for mouse κ-chains (Hood *et al.*, 1970) on the basis of a study of the *N*-terminal 23 residues of 22 light chains of the Balb/c strain, but these basic sequences in our opinion, are not so well defined. This situation parallels the pattern observed in human λ-chains, where assignment of basic sequences on the basis of small stretches of *N*-terminal sequence is somewhat suspect (Milstein & Pink, 1970) and can confidently be achieved only by using more extensive sequence information.

Complete sequences of mouse κ-chains have not been reported, but extensive sequences on two mouse κ-chains have been reported by Gray *et al.* (1967).

In the present paper we report the complete amino acid sequence of a mouse κ light chain, derived from the plasmacytoma, MOPC 21 (IgG1*).

*Abbreviations: IgG, immunoglobulin G; CmCys, $CmCys(O_2)$, $Cys(O_3H)$ and $Met(O_2)$ (in amino acid sequences and tables), *S*-carboxymethylcysteine, *S*-carboxymethylcysteine sulphone, cysteic acid and methionine sulphone respectively.

A preliminary account of this study has been reported (Milstein & Svasti, 1971).

Materials and Methods

Materials

The plasmacytoma, MOPC 21, from Dr. M. Potter's collection, was maintained by transplantation in Balb/c mice. Serum was kindly supplied by Dr. A. J. Munro and Mr. P. Wright.

Iodo[2-^{14}C]acetic acid (10–20 mCi/mmol) was obtained from The Radiochemical Centre, Amersham, Bucks., U.K.

Trypsin, chymotrypsin and pepsin were from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Leucine aminopeptidase was obtained from C. F. Boehringer und Soehne G.m.b.H., Mannheim, Germany. Thermolysin was a gift from Dr. R. P. Ambler.

Reagents were generally of analytical grade. HCl was obtained as a 34–36% solution (Aristar grade; BDH Chemicals Ltd., Poole, Dorset, U.K.) and was diluted with an equal vol. of water. Phenol (1 mg/ml) was added to prevent destruction of tyrosine (Sanger & Thompson, 1963). Reagents for the Edman degradation cycle were redistilled as described by Edman & Begg (1967).

Methods

Purification of myeloma protein and isolation of light chains. This was done as described by Svasti & Milstein (1972).

Total reduction and alkylation with iodo[^{14}C]acetic acid. This was done on separated light chains as described by Milstein (1968). The solution of totally

reduced and carboxymethylated light chains was exhaustively dialysed against 1% (w/v) NH_4HCO_3 and then was digested with trypsin or chymotrypsin.

High-voltage paper electrophoresis of peptides. This was done as described by Ambler (1963) or Milstein (1966) for 45–120min. Mobilities (*m*) at pH6.5 are expressed as fractions of the distance between the aspartic acid and ε-Dnp-lysine spots (Ambler, 1963; Offord, 1966). Amide residues were assigned on the basis of electrophoretic mobility at pH6.5 (Offord, 1966). Mobilities at pH3.5 are expressed as fractions of the distance between the glycylalanine and ε-Dnp-lysine spots, and mobilities at pH2.0 are relative to valine.

Paper chromatography was performed with a descending butan-1-ol–acetic acid–water–pyridine system (15:3:12:10, by vol.) (solvent BAWP) (Waley & Watson, 1953). R_F values are expressed relative to valine.

Enzyme digests of carboxymethylated protein and fractionation of digests. Preliminary tryptic and chymotryptic digests were separately performed on 50mg of protein at an enzyme/substrate ratio of 1:100 (w/w) for 4.5h at 37°C in 1% (w/v) NH_4HCO_3. Each digest was subjected to electrophoresis at pH6.5 on Whatman 3MM paper with a loading of 1.2mg/cm, and 5cm-wide strips of the acidic and basic sections of the chromatogram (but not including the neutral peptides) were sewn on to a new sheet of paper. Two-dimensional 'fingerprints' of the acidic and basic peptides were obtained by chromatography in solvent BAWP, run at right angles to the original run. The neutral tryptic peptides, including the origin ('core') peptides, were run at pH8.9 in the same direction as the pH6.5 run, and then further fractionated by chromatography in solvent BAWP run at right angles to the previous runs (preparative 'fingerprints', see Fig. 1). The chymotryptic neutral peptides were fractionated by electrophoresis at pH2.0, followed by chromatography in solvent BAWP. The papers were radioautographed; the radioactive peptides were cut out and the rest of the paper was stained with 0.03% (w/v) ninhydrin in acetone (Anfinsen *et al.*, 1959) to locate the other peptides. Each peptide was eluted with water (except for peptides T7, T9 and T10, which were eluted with 0.1 M-NH_3), and hydrolysed with 6 M-HCl for analysis.

In addition, analytical 'fingerprints' (carried out at pH6.5 and then at pH3.5 for acidic and basic peptides, and at pH6.5 and then at pH2.0 and then in solvent BAWP, for neutral tryptic peptides) were also made and stained with cadmium–ninhydrin reagent (Heilmann *et al.*, 1957) to determine the best purification procedure for each peptide.

It was apparent that not all the tryptic peptides could be purified in good yield simply by paper methods. Thus another tryptic digest was performed on 200mg of carboxymethylated protein under the conditions as described above; it was stopped by addition of 2mg of soya-bean trypsin inhibitor and then was applied to a column (4cm×120cm) of Sephadex G-50 (fine grade), equilibrated with 1% (w/v) NH_4HCO_3. Radioactivity was monitored in a Nuclear-Chicago flow scintillation counter. Samples

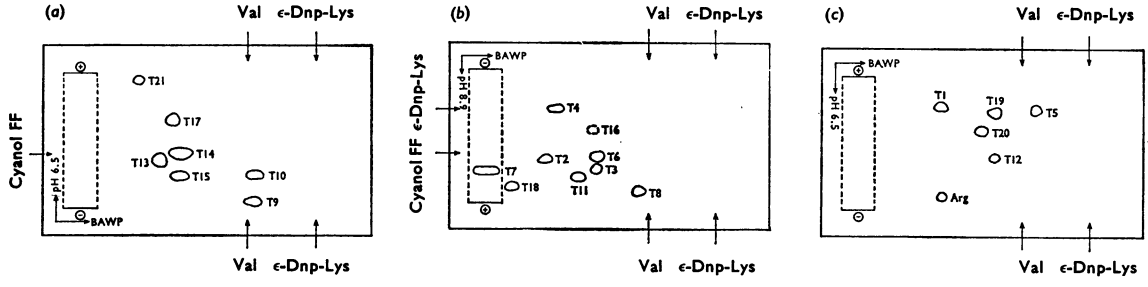

Fig. 1. *Preparative 'fingerprints' of a tryptic digest of reduced and* [^{14}*C*]*carboxymethylated MOPC* 21 *light chain*

The digest was subjected to electrophoresis at pH6.5, and the paper was divided into three sections, acidic peptides, neutral peptides, including the origin ('core') peptides, and basic peptides. Acidic peptides (*a*) and basic peptides (*c*) were run in chromatography (solvent BAWP) at right angles to the original run. Neutral and 'core' peptides (*b*) were subjected to electrophoresis at pH8.9 in the same direction as the pH6.5 run, followed by chromatography (solvent BAWP) at right angles to the previous runs. [^{14}C]Carboxymethylcysteine peptides were identified by radioautography, tryptophan peptides by fluorescence, and the remainder of the paper was stained with 0.03% ninhydrin in acetone. Arrows indicate the positions of markers: ε-Dnp-Lys, Cyanol FF and valine. Boxes bounded by broken lines show the positions of stitching. Two peptides (T1 and T15) were also isolated in other forms, which differed in mobility at pH6.5. However, as these forms had probably arisen as a result of deamidation, they are not shown in Fig. 1. Peptide T16 is shown in broken lines as it could not be isolated from all digests. It was suspected that the solubility of T16 in chromatography (solvent BAWP) may depend on the order in which the pH8.9 electrophoresis and chromatography steps are performed.

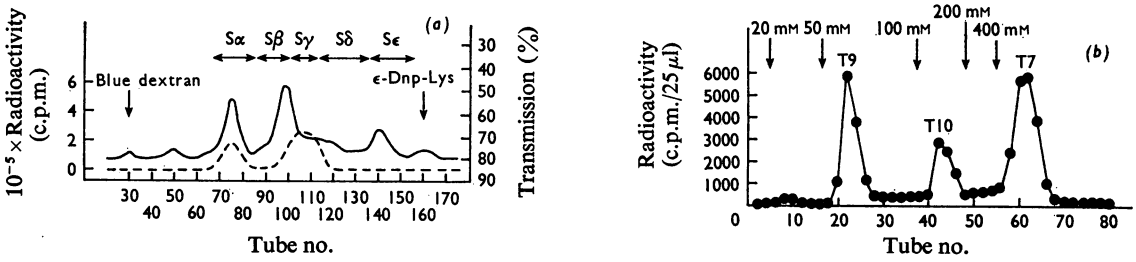

Fig. 2. *Purification of the large [^{14}C]carboxymethylcysteine peptides (T7, T9 and T10) from a tryptic digest of reduced and alkylated MOPC 21 light chain*

(*a*) Fractionation of the tryptic digest in a column (4cm × 120cm) of Sephadex G-50 (fine grade), equilibrated with 1% (w/v) NH_4HCO_3; 4ml fractions were taken. ——, E_{254} in arbitrary units obtained in an LKB Uvicord spectrophotometer; – – –, radioactivity in arbitrary units obtained in a Nuclear–Chicago 6770 Chromacell flow scintillation counter. The higher-molecular-weight material (pool Sα) was freeze-dried, was dissolved in 5ml of dilute NH_3 and was applied to a column (1cm × 10cm) of DEAE-cellulose, equilibrated with 20mM-ammonium acetate, pH8.1. (*b*) Purification of large tryptic peptides from Sephadex pool Sα by chromatography in a column of DEAE-cellulose, equilibrated with 20mM-ammonium acetate, pH8.1. Arrows indicate the ammonium acetate concentration; 2.5ml fractions were taken and 25μl samples were removed to determine radioactivity in a scintillation counter.

were removed, applied to filter paper, and stained for the presence of tryptophan, as described by Smith (1953). Fractions were pooled as shown in Fig. 2(*a*) and freeze-dried.

The higher-molecular-weight fractions (Sα and Sβ) were separately purified by chromatography in a column (1cm × 8cm) of DEAE-cellulose (Whatman DE-52), equilibrated with 20mM-ammonium acetate, pH8.1. Fraction Sα was purified by stepwise elution with increasing concentrations of ammonium acetate, pH8.1 (Fig. 2*b*) to yield the radioactive peptides T10, T9 and T7. A gradient of ammonium acetate (pH8.1, 20–500mM) was used to purify fraction Sβ. Peptide T4, characterized by strong u.v. absorption and by the presence of tryptophan, was eluted pure at approx. 50mM. Sephadex fractions Sγ, Sδ and Sε were not further characterized.

Enzyme digestion of peptides. Thermolysin digests of peptides were performed in 1% (w/v) ammonium acetate–0.005M-$CaCl_2$ buffer, pH8.5, at 45° or 60°C for 2–5h with 10–20μg of enzyme. Leucine aminopeptidase digests were performed in 5mM-tris–HCl–5mM-$MgCl_2$ buffer, pH8.2, at 37°C with 10–20μg of enzyme. Amino acids, liberated at different times, were determined in a Beckman 120C amino acid analyser. In addition, residual peptides were sometimes characterized as described in the text. Pronase and pepsin digests were performed by using 10–20μg of enzyme for 2–4h at 37°C, in 1% (w/v) ammonium acetate buffer, pH7, and in 5% (v/v) formic acid respectively.

Detection of non-radioactive peptides on paper, acid hydrolysis and amino acid analysis were carried out as described by Milstein (1966). Peptides were eluted with water, 0.1M-NH_3 or 0.5% acetic acid.

Determination of tryptophan. Tryptophan was determined quantitatively by the method of Spies & Chambers (1948), as described by Ambler & Brown (1967).

Performic acid oxidation. Oxidation of the chain (for amino acid analysis) was done as described by Hirs (1967), whereas oxidation of peptides was performed as described by Brown & Hartley (1966).

Dansyl–Edman procedure. This was done as described by Gray (1967*b*). Dns-amino acids were identified by two-dimensional t.l.c. (Woods & Wang, 1967). A third solvent (solvent IV of Crowshaw *et al.*, 1967) was used to separate aspartic acid from glutamic acid, and serine from threonine. *N*-Terminal [^{14}C]carboxymethylcysteine was identified by the loss of radioactivity after the next Edman degradation cycle (Svasti & Milstein, 1972). The last amino acid in each peptide was identified as a free amino acid either by identification of the Dns derivative without prior hydrolysis (Pink *et al.*, 1970) or by amino acid analysis (Milstein, 1966). *N*-Terminal analysis of the whole chain was done by the method of Gray (1967*a*), as described by Pink *et al.* (1970).

The results of sequence determination are shown under the peptide on which the experiments were performed. Arrows (→) indicate the results of the dansyl–Edman procedure, double arrows (⇒) indicate the release of a free *C*-terminal amino acid; broken arrows (⇢) indicate a dubious or negative result. Half arrows (⇀) indicate results of leucine aminopeptidase digestion.

Table 1. *Amino acid composition of MOPC 21 light chain*

Compositions were determined from four hydrolyses at 24, 48, 72 and 96h of performic acid-oxidized protein and one 24h hydrolysis of untreated protein. Optimum values for serine and threonine were calculated by extrapolation to zero time; those for valine and isoleucine were calculated by extrapolation to infinite time; and those for other amino acids were calculated by taking an average of the five hydrolyses.

	Untreated protein	Oxidized protein				Optimum	Value calculated
Time of hydrolysis ...	24h	24h	48h	72h	96h	value	from sequence
Lys	13.2	12.1	12.9	12.7	13.6	12.9	13
His	3.1	2.8	3.0	3.0	3.2	3.0	3
Arg	6.7	6.5	6.8	6.7	6.9	6.7	7
$Cys(O_3H)$*	—	4.4	4.6	4.3	4.5	4.5	5
Asp	20.6	22.9	22.4	21.8	22.1	22.0	22
$Met(O_2)$	—	4.0	3.9	4.0	4.1	4.0	4
Thr	20.0	19.4	18.3	17.6	15.9	22.2	22
Ser	26.2	26.7	23.9	20.2	19.6	28.6	29
Glu	19.4	20.4	19.8	19.4	20.2	19.8	20
Pro	11.1	9.7	10.4	9.8	10.4	10.3	10
Gly	14.8	13.4	13.2	14.3	13.5	13.8	14
Ala	10.8	10.0	9.4	9.9	10.0	10.0	10
Val	12.9	13.6	14.0	14.0	14.2	14.1	14
Ile	7.4	7.2	7.7	7.8	8.0	8.0	8
Leu	11.7	10.8	10.8	10.9	10.5	10.9	11
Tyr	12.0	10.5	10.1	10.2	9.8	12.0	12
Phe	7.2	6.8	6.7	7.0	6.7	6.9	7
Trp†	3.0					3.0	3

* Carboxymethylcysteine sulphone obtained from the *C*-terminal half-cystine residue, when the light-chain was separated from heavy-chain after partial reduction and alkylation, is included.

† Tryptophan was determined as described in the Materials and Methods section.

Table 2. *Amino acid compositions (residues/mol of peptide), yields (expressed as mol of peptide*

Compositions and yields of all peptides, except for T4, T7, T9 and T10, are determined from the preparative 'fingerprints' shown in Fig. 1. T4, T7, the Materials and Methods section). The best method for purification is also shown; paper electrophoresis is indicated by the pH used, and paper

Peptide ...	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
Purification procedure ...	6.5, BAWP	6.5, BAWP, 2.0	6.5, BAWP, 2.0	G-50, DEAE	6.5, BAWP	6.5, BAWP, 8.9	G-50, DEAE	6.5, BAWP, 2.0	G-50, DEAE	G-50, DEAE	6.5, BAWP, 2.0
Lys	1.0		0.8	1.9			1.0	0.8	0.9	0.8	0.8
His							1.0				
Arg		1.0			1.0	1.0			0.9		
CmCys			0.6				0.8		0.8	1.0	
Asp	0.9			1.0	1.0	1.0	3.1		2.9	2.8	2.0
Thr	1.1		1.8	1.0		1.0	5.5		1.9	1.9	
Ser	1.1	2.8		2.9	0.9		4.5		4.6	5.2	
Glu	1.1	1.1		5.0			2.9	1.1	1.8	1.8	
Pro	1.0			2.1		1.0	1.0		4.1	3.7	
Gly		1.1		0.2	1.1	1.1	6.9		2.0	2.1	
Ala				1.1	1.1		2.8		4.2	3.9	
Val	0.8	1.1	0.8	2.9*		1.0	1.1		2.5	2.8	1.1
Met	1.0	2.0									
Ile	0.75				0.9		1.0	1.0	1.0	1.0	1.1
Leu			1.1		1.9		2.1	0.9	2.0	2.1	
Tyr				2.0	0.8	0.6	4.1		1.1	1.0	
Phe							3.0		3.0	2.8	
Trp				+							
N-Terminus	Asx	Ser	Val	Ala	Leu	Tyr	Phe	Leu	Arg	Ala	Asx
m (pH6.5)	−0.19	0	0	0	−0.25	0	Insol.	0	0.05	0.18	0
% Yield	9.2	18.5	16.2	28.3	21.8	21.7	50	25.2	35	15	17.9

* Value after 96h hydrolysis; after 20h the value obtained was 2.3.

Results

The amino acid composition of MOPC 21 light chain was determined by extrapolation from four hydrolyses at 24, 48, 72 and 96h of performic acid-oxidized protein and one 24h hydrolysis of untreated protein. The results are shown in Table 1. The chain was shown to contain an aspartic acid (or the amide) as the *N*-terminal residue.

Tryptic peptides of [^{14}C]carboxymethylated light chain

Most of the tryptic peptides could be simply purified by paper electrophoresis and paper chromatography. The larger tryptic peptides (T4, T7, T9 and T10) were purified in better yields by Sephadex G-50 gel filtration followed by DEAE-cellulose chromatography. The amino acid composition of each tryptic peptide is shown in Table 2, together with the best method found for its purification.

The value for the yield of peptides was obtained from preparative 'fingerprints' (Fig. 1), except for peptides T4, T7, T9 and T10, when it was obtained by the Sephadex G-50–DEAE-cellulose purification procedure.

Peptide T1. This was isolated with two mobilities at pH6.5 (−0.19 and 0), suggesting that the basic peptide was the original form and that the neutral peptide arose from it by deamidation of the *N*-terminal asparagine residue. This view is supported by the mobilities of the chymotryptic peptides, C1 and C2 (see below, Table 9, Fig. 7). The sequence (Fig. 3) of the peptide was established by the dansyl–Edman procedure.

Peptides T2 and T3. These are described in Fig. 3.

Peptide T4. This peptide was best purified from fraction Sβ (Fig. 2*a*) by DEAE-cellulose chromatography (see the Materials and Methods section). It was neutral on paper electrophoresis at pH6.5, indicating the presence of two acidic groups. Its sequence was determined by the dansyl–Edman procedure (Fig. 3) and confirmed by isolation of the thermolysin derivatives, T4HA, T4HNF, T4HN1, T4HB (Table 3), and the chymotryptic derivatives, T4CA, T4CNF and T4CB (Table 4). Peptide T4HA had a mobility of 0.45, corresponding to a tetrapeptide with one net charge. The Dns derivative of its last residue was identified as Dns-Asn or Dns-Gln, whereas after hydrolysis Dns-Asp was obtained. This established the sequence Ala-Ser-Glu-Asn.

Peptide T4HB had a mobility of −0.25, indicating the presence of only one glutamic acid residue. It was digested with leucine aminopeptidase, and a time-course of liberation of free amino acids was determined with an amino acid analyser, under conditions where glutamine was well resolved from serine (Fig. 4). It is clear that initially one tyrosine and then two glutamine residues were released, indicating the *N*-terminal sequence Tyr-Gln-Gln. The position of the third glutamine residue is not clear, but glutamic acid is clearly released before lysine, in direct contradiction with the results obtained by the dansyl–Edman procedure on peptides, T4 and T4CB. The explanation for this is as follows: leucine aminopeptidase begins

---

recovered/mol of protein used) and N-termini of the tryptic peptides of MOPC 21 light chain

T9 and T10 peptides are determined from the Sephadex G-50–DEAE-cellulose purification procedure (indicated by G-50, DEAE and described in chromatography is indicated by BAWP. Electrophoretic mobilities at pH6.5 (*m*) are expressed relative to aspartic acid.

Peptide ...	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21
Purification procedure ...	6.5, BAWP	6.5, BAWP, 3.5	6.5, BAWP	6.5, BAWP	6.5, 8.9, BAWP	6.5, BAWP	6.5, BAWP, 2.0	6.5, BAWP	6.5, BAWP	6.5, 3.5
Lys	+		0.9	0.8	0.8		0.9	0.9		
His							1.5			
Arg		0.9	0.8			0.9			0.9	
CmCys							0.7			0.7
Asp		1.1	5.0	4.0	0.9	0.8	1.1		1.1	1.0
Thr			1.0	0.9	4.0		2.1	2.0		
Ser		0.9	2.8	1.8	3.9		1.0	2.1	0.8	
Glu		1.0	3.1	1.9		2.1	1.1			1.0
Pro								1.1		
Gly		1.1	2.3	1.0						
Ala							0.9			
Val			1.2	0.9				0.8		
Met					1.0					
Ile		0.9	0.9					0.7		
Leu			1.0	0.9	2.1					
Tyr					1.0	1.1	0.8			
Phe									1.0	
Trp	+		+	+						
N-Terminus	—	Ile	Ile	Glx	Asx	Asx	His	Thr	Ser	Asx
m (pH6.5)	−0.52	0.33	0.36	0.20	0	0.65	−0.05	−0.28	−0.40	0.95
% Yield	12.8	22	11.5	7.5	11.9	12.5	9.3	26.5	24	20

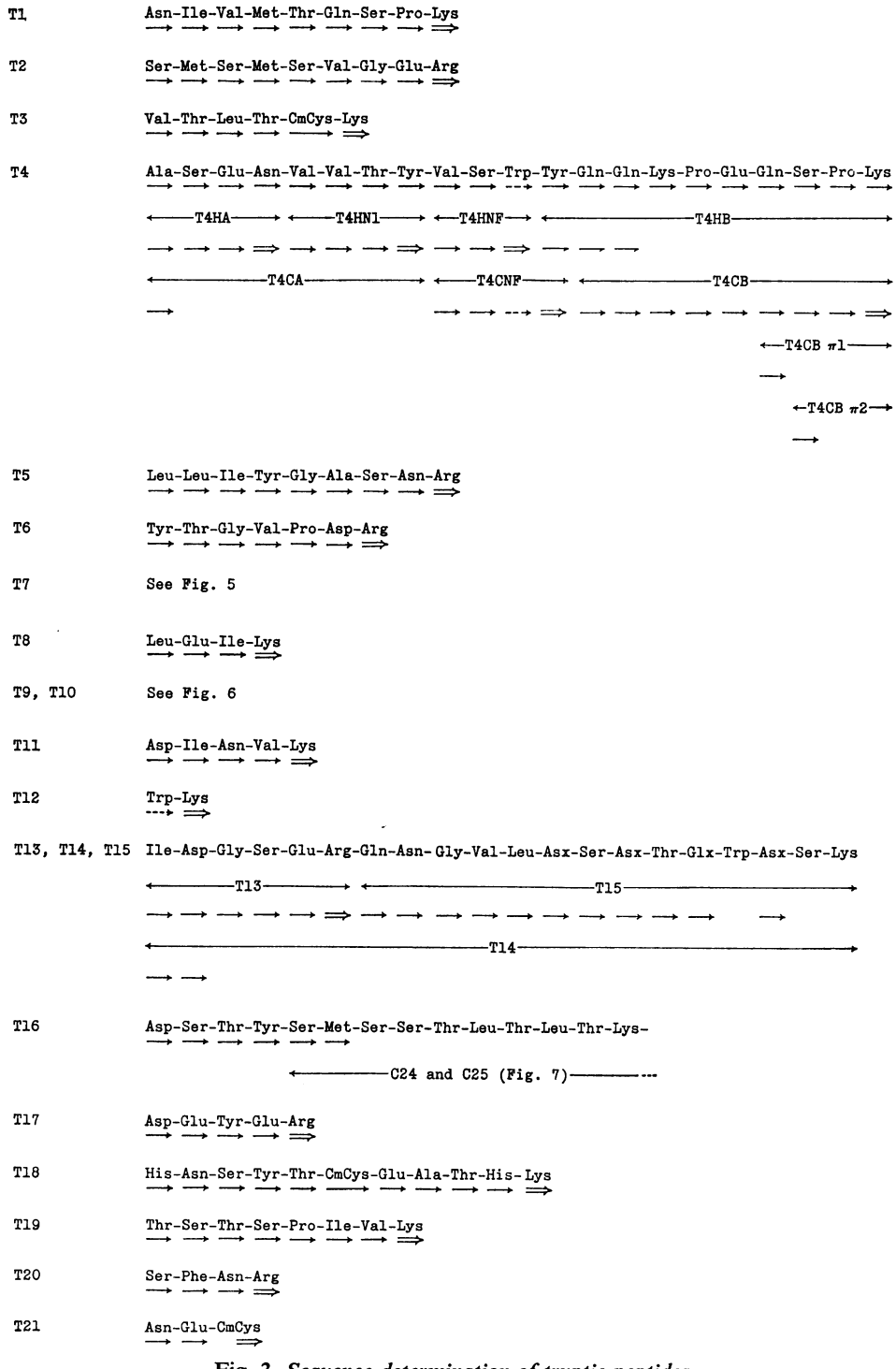

Fig. 3. *Sequence determination of tryptic peptides*

Amino acid compositions are shown in Table 2 (see also Tables 3 and 4 and Fig. 4 for peptide T4). For details see the text.

Table 3. *Peptides isolated from peptide T4 after digestion with thermolysin*

Purification was by electrophoresis at the indicated pH or by chromatography (BAWP). Dansyl–Edman results are shown in Fig. 3. Neutral peptides were eluted after staining with 0.03% ninhydrin. Relative yields of the peptides were determined by amino acid analysis.

Peptide	Purification procedure	*m* (pH6.5)	Relative yield	Amino acid composition (residues/mol)
T4HA	6.5	0.45	+++	Asp (1.0), Ser (0.9) ,Glu (1.0), Ala (1.0)
T4HN1	6.5, 8.9, BAWP	Neutral	++	Thr (1.0), Val (1.5), Tyr (1.0)
T4HNF	6.5, 8.9, BAWP	Neutral	++	Val (1.0), Ser (0.9), Trp (+)
T4HB	6.5	−0.25	+++	Tyr (1.0), Ser (1.0), Glu (4.1), Pro (2.0), Lys (2.0)

Table 4. *Peptides isolated from peptide T4 after digestion with chymotrypsin and peptides isolated from a Pronase digest of peptide T4CB*

Purification was by electrophoresis at pH6.5. Dansyl–Edman results are shown in Fig. 3. Relative yields of the peptides were determined by amino acid analysis. Peptides T4CA, T4CNF and T4CB were isolated after digestion of peptide T4 with chymotrypsin. Peptides T4CBπ1 and T4CBπ2 were isolated from a Pronase digest of peptide T4CB.

Peptide	Purification procedure	*m* (pH6.5)	Relative yield	Amino acid composition (residues/mol)
T4CA	6.5	0.28	+++	Asp (1.1), Thr (1.0), Ser (1.0), Glu (1.1), Ala (0.9), Val (1.4), Tyr (1.0)
T4CNF	6.5	−0.03	+++	Ser (1.0), Val (1.0), Tyr (0.8), Trp (+)
T4CB	6.5	−0.28	+++	Ser (0.9), Glu (3.9), Pro (2.1), Lys (1.8)
T4CBπ1	6.5	−0.38	+	Ser (1.0), Glu (1.0), Pro (1.0), Lys (1.0)
T4CBπ2	6.5	−0.48	+	Ser (1.0), Pro (1.0), Lys (1.0)

stepwise degradation from the *N*-terminal residue but stops when an X–Pro bond is reached (in this case Lys-Pro-Glx-Glx...). Then the enzyme (or, more likely, a contaminating protease) splits the Pro–Glx bond to yield the dipeptide Lys-Pro and the residual peptide, which is further degraded until the peptide Ser-Pro-Lys is obtained. Cleavage now occurs at the Pro–Lys bond to give Ser-Pro and free lysine. This explanation is supported by the fact that no free serine or proline was obtained and, more important, by the isolation of the dipeptide:

$$\text{Lys-Pro } (m = -0.61;\ \text{Lys } 1.0,\ \text{Pro } 1.0) \xrightarrow{} \Rightarrow$$

in addition to free lysine after aminopeptidase digestion was complete (6h). The yield of the dipeptide Lys-Pro was approximately the same as that of free lysine and their combined yields were too high to be derived from the same lysine residue. Release of X-Pro dipeptides by aminopeptidase M have been reported previously (Plummer, 1969; see also Delange & Smith, 1971).

Peptides T4CBπ1 and T4CBπ2 were isolated from a Pronase digest of the chymotryptic derivative T4CB (Table 4). Since peptide T4CBπ1 was basic, the Glx residue in that peptide must be glutamine, but

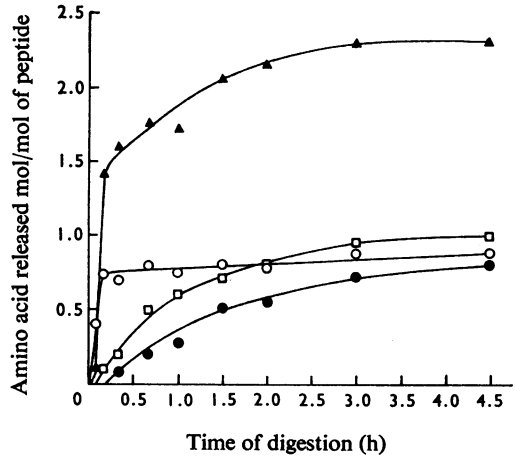

Fig. 4. *Time-course of release of amino acids by digestion of peptide T4 with leucine aminopeptidase*

For experimental details see the text. Samples were taken and the release of free amino acids was determined in an amino acid analyser: o, tyrosine; ▲, glutamine; □, glutamic acid; ●, lysine.

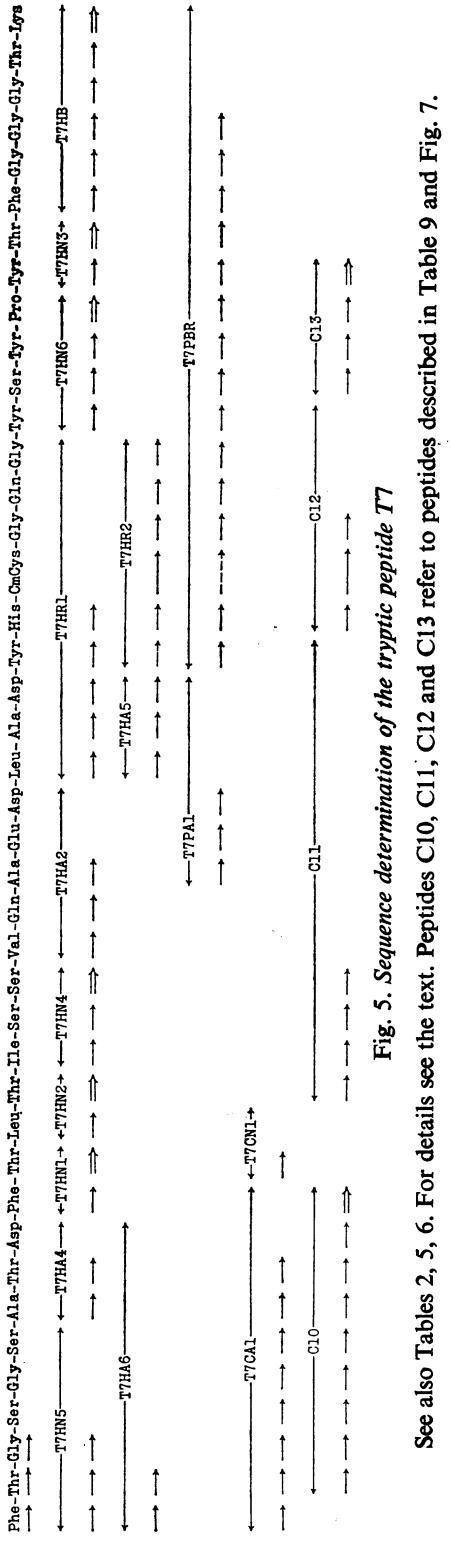

Fig. 5. *Sequence determination of the tryptic peptide T7*

See also Tables 2, 5, 6. For details see the text. Peptides C10, C11, C12 and C13 refer to peptides described in Table 9 and Fig. 7.

peptide T4CB (and peptide T4HB) must include one acidic residue. The final assignment of amide groups in peptide T4HB is therefore Tyr-Gln-Gln-Lys-Pro-Glu-Gln-Ser-Pro-Lys.

Peptides T5 and T6. Sequences are shown in Fig. 3. Amide groups are assigned from their mobilities at pH6.5.

Peptide T7. This peptide was insoluble at pH6.5 and was isolated as a radioactive peptide by Sephadex G-50 and DEAE-cellulose chromatography. It was digested with thermolysin (Fig. 5, Table 5). Peptide T7HA2 had a mobility of 0.66, indicating the presence of two acidic groups. After two Edman degradation cycles, the residual peptide T7HA2A3 had a mobility of 0.95, indicating the amide assignment as Val-Gln-Ala-Glu-Asp. Peptide T7HR2 was almost neutral at pH6.5 indicating the position of an amide in the sequence Tyr-His-CmCys-Gly-Gln-Gly. Overlapping peptides for use in establishing the complete sequence of peptide T7 were obtained from a peptic and from a chymotryptic digest of peptide T7. A number of peptides were isolated from these digests, but only those providing essential overlaps are shown in Table 6 and Fig. 5. The overlap between peptides T7HN2 and T7HN4 was provided by the peptide C11 (Table 9, Fig. 7), isolated from a chymotryptic digest of intact light chain.

Peptide T8. This was neutral, indicating that it contained glutamic acid rather than the amide. Fig. 3 shows the sequence.

Peptides T9 and T10. These radioactive peptides were only partially soluble at pH6.5 and were purified in excellent yields by Sephadex G-50 gel filtration, followed by DEAE-cellulose chromatography (Fig. 2*a* and 2*b*). The two very-well-resolved peptides differed only in the presence of an arginine residue at the *N*-terminus of peptide T9, owing to an incomplete tryptic split (Fig. 8). Free arginine was also isolated from the digest. Peptide T9 was digested with thermolysin to give the peptides shown in Table 7 and Fig. 6. The peptide T9HA had a mobility of 0.29, indicating that one of the two glutamic acid residues was in the amide form. The Dns derivative of the *C*-terminal residue was Dns-Gln or Dns-Asn, giving Dns-Glu after acid hydrolysis. The other amide assignments were deduced from electrophoretic mobilities at pH6.5. Overlapping peptides were obtained by digestion of peptide T9 with chymotrypsin (Table 8, Fig. 6).

Peptide T11. This peptide was neutral at pH6.5, indicating the presence of one aspartic acid and one asparagine residue. The sequence was established by the dansyl–Edman procedure (Fig. 3). The major spots on electrophoresis at pH6.5 after one and two cycles of Edman degradation were basic, indicating the sequence Asp-Ile-Asn-Val-Lys.

Peptide T12. The mobility of this peptide was consistent with that of a dipeptide with one charge.

Table 5. *Peptides isolated from a thermolysin digest of the tryptic peptide T7*

Purification was by paper electrophoresis at the indicated pH and chromatography (BAWP). Dansyl–Edman results are shown in Fig. 5. Neutral peptides were eluted after staining with 0.03% ninhydrin. Relative yields of the peptides were determined by amino acid analysis.

Peptide	Purification procedure	*m* (pH 6.5)	Relative yield	Amino acid composition (residues/mol)
T7HR1	6.5, 3.5	0.28	++	His (1.1), CmCys (0.5), Asp (1.2), Glu (1.0), Gly (2.1), Ala (0.9), Tyr (0.7), Leu (0.9)
T7HR2	6.5, 3.5	0.07	+++	His (1.1), CmCys (+), Glu (1.1), Gly (1.9), Tyr (1.0)
T7HA2	6.5, 2.0	0.66	+++	Asp (0.8), Glu (2.2), Ala (0.9), Val (1.0)
T7HA4	6.5, 2.0	0.54	+++	Asp (1.0), Thr (1.0), Ala (0.9)
T7HA5	6.5, 2.0	0.49	+++	Asp (1.1), Ala (1.0), Leu (0.9)
T7HA6	6.5, 3.5	0.24	+	Asp (1.1), Thr (2.0), Ser (2.0), Gly (2.1), Ala (1.1), Phe (1.0)
T7HN1	6.5, 2.0, BAWP	0	+++	Thr (1.0), Phe (0.9)
T7HN2	6.5, 2.0, BAWP	0	+++	Thr (1.0), Leu (0.9)
T7HN3	6.5, 2.0, BAWP	0	+++	Thr (1.0), Tyr (0.8)
T7HN4	6.5, 2.0, BAWP	0	++	Ser (2.0), Ile (1.0)
T7HN5	6.5, 2.0, BAWP	0	++	Thr (1.0), Ser (2.0), Gly (2.2), Phe (0.8)
T7HN6	6.5, 2.0, BAWP	0	++	Ser (1.1), Pro (1.0), Tyr (1.6)
T7HB	6.5, 2.0	−0.38	+++	Lys (1.0), Thr (1.1), Gly (3.0), Phe (0.9)

Table 6. *Peptides isolated from digests of tryptic peptide T7 with pepsin (peptide names beginning with T7P) or chymotrypsin (peptide names beginning with T7C)*

Only peptides providing essential sequence data or overlaps are shown. Purification was by electrophoresis at the indicated pH or by chromatography (BAWP). Dansyl–Edman results are shown in Fig. 5. Peptide T7CN1 was eluted after staining with 0.03% ninhydrin. Relative yields of the peptides were determined by amino acid analysis.

Peptide	Purification procedure	*m* (pH 6.5)	Relative yield	Amino acid composition (residues/mol)
T7PA1	6.5	0.79	++	Asp (2.0), Glu (1.2), Ala (1.9), Leu (1.0)
T7PBR	6.5	−0.13	++	Lys (0.9), His (0.9), CmCys (+), Thr (1.8), Ser (0.9), Glu (1.1), Pro (1.0), Gly (4.9), Tyr (3.7), Phe (1.0)
T7CA1	6.5	0.24	+++	Asp (1.1), Thr (2.0), Ser (1.9), Gly (2.0), Ala (1.1), Phe (1.9)
T7CN1	6.5, 2.0, BAWP	0	++	Thr (0.9), Leu (1.0)

The *N*-terminal residue could not be definitely identified as Dns-Trp, but free lysine was obtained after one cycle of Edman degradation (Fig. 3).

Peptides T13, T14, T15. The amino acid composition of peptide T14 was the sum of the compositions of peptides T13 and T15 (Table 2). This, together with the studies on the chymotryptic peptide C22 (Table 9, Fig. 7) and the dansyl–Edman degradation, showed that peptide T14 was the result of an incomplete tryptic split at the arginine residue of peptide T13 (see Fig. 3). The mobility of peptide T14 indicated two net acidic charges, and peptides T13 and T15 had electrophoretic mobilities (Table 2) indicating one net acidic charge in each. Taken together with the amino acid compositions (Table 2), the results indicate that peptide T13 had no amide groups, whereas peptide T15 had two acidic and four amide groups. Two of these amide groups could be assigned as Gln-Asn-Gly ..., from the neutral chymotryptic peptide C22 (Table 9, Fig. 7), which can only include two acid residues both present in its peptide-T13 moiety. However, it was not possible to assign the remaining amide residues, because we found that different preparations gave erratic mobilities. Peptide T15, isolated (in better yields) from Sephadex fraction Sβ, gave several derivatives with mobilities 0.65,

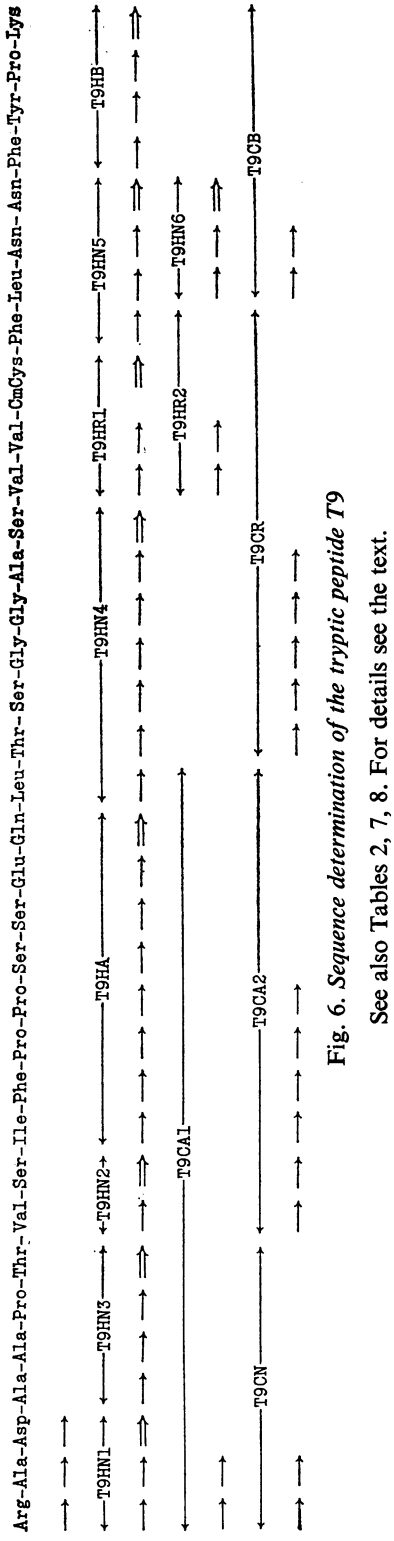

Fig. 6. *Sequence determination of the tryptic peptide T9*

See also Tables 2, 7, 8. For details see the text.

0.52, 0.38, corresponding to four, three and two net charges respectively, probably arising as a result of deamidation. In addition, the peptide with mobility 0.65 only yielded ε-Dns-Lys on reaction with dansyl chloride, and therefore lacked a free α-amino group. This indicates that glutamine had cyclized to form pyrrolid-2-one-5-carboxylic acid. In a previous paper (Milstein & Svasti, 1971), we reported a sequence with the amide groups identified from such preparations, but pointed out that some of the acidic groups in this peptide were likely to be the result of deamidation. As preparation of peptide T15 in good yields without deamidation derivatives has been clearly erratic, we have chosen in the present paper to leave the suspected residues as undetermined amides.

Peptide T16. This peptide was neutral, indicating that the Asx residue was aspartic acid. Peptide T16 could not be isolated from all digests, probably because it may only have been soluble in solvent BAWP if it had previously been subjected to electrophoresis at pH8.9. The *N*-terminal sequence was determined by the dansyl–Edman procedure (Fig. 3). The final sequence was deduced from studies on chymotryptic peptides C24 and C25 (Fig. 7).

Peptide T17. This peptide had a mobility of 0.65 at pH6.5 indicating that three acidic groups were present. Fig. 3 shows the sequence.

Peptide T18. Fig. 3 shows the sequence. The amino acid analysis shows one glutamic acid and one aspartic acid residue in addition to carboxymethylcysteine; the peptide also contains two residues of histidine and one of lysine. As this peptide was slightly basic at pH6.5 (Table 2), it cannot include more than one (and probably not less than one) acidic residue in addition to carboxymethylcysteine. The mobility is, in fact, consistent with the assignment of only one amide residue to either the aspartic acid or glutamic acid residues and of approx. 0.5 charge for each of the histidine residues. The mobility at pH8.9 indicates two acidic residues (carboxymethylcysteine and either aspartic acid or glutamic acid). This is better shown by the greater acidity of this peptide than that of peptide T3 (Fig. 1*b*), which has one acidic residue and about half the molecular weight of peptide T18. The assignment of the acidic charge to the glutamic acid comes from studies on two chymotryptic peptides (C26 and C27) discussed below (Fig. 7 and Table 9), and is confirmed by the isolation of a thermolytic derivative T18HR1 ($m = 0.75$; CmCys positive, Thr 0.9, Glu 1.2, Tyr 0.9), which must include two net acidic charges.

Peptides T19 and T20. The sequences of these peptides were established by the dansyl–Edman procedure (Fig. 3).

Peptide T21. This peptide was radioactive and had a mobility of 0.95, corresponding to two net charges. After one Edman degradation cycle the residual

Table 7. *Peptides isolated from a thermolysin digest of peptide T9*

Purification was by electrophoresis at the indicated pH and by chromatography (BAWP). Dansyl–Edman results are shown in Fig. 6. Neutral peptides were eluted after staining with 0.03% ninhydrin. Relative yields of the peptides were determined by amino acid analysis.

Peptide	Purification procedure	*m* (pH 6.5)	Relative yield	Amino acid composition (residues/mol)
T9HR1	6.5, 3.5	0.53	++	CmCys (0.6), Val (2.0)
T9HR2	6.5, 3.5	0.43	+	CmCys (0.5), Val (1.7), Phe (1.0)
T9HA	6.5, 2.0	0.29	+++	Ser (2.0), Glu (2.1), Pro (2.0), Ile (1.0), Phe (1.1)
T9HN1	6.5, 2.0, BAWP	0	++	Arg (0.8), Asp (1.0), Ala (1.1)
T9HN2	6.5, 2.0, BAWP	0	+++	Ser (1.0), Val (1.0)
T9HN3	6.5, 2.0, BAWP	0	+++	Thr (1.0), Pro (1.0), Ala (1.9)
T9HN4	6.5, 2.0, BAWP	0	+++	Thr (1.0), Ser (1.8), Gly (2.0), Ala (1.0), Leu (0.8)
T9HN5	6.5, 2.0, BAWP	0	+	Asp (2.0), Leu (1.0), Phe (0.9)
T9HN6	6.5, 2.0, BAWP	0	++	Asp (2.1), Leu (0.9)
T9HB	6.5	–0.38	++	Lys (1.0), Pro (0.9), Tyr (1.0), Phe (0.9)

Table 8. *Peptides isolated from a chymotryptic digest of peptide T9*

Purification was by electrophoresis at the indicated pH. Dansyl–Edman results are shown in Fig. 6. Relative yields of the peptides were determined by amino acid analysis.

Peptide	Purification procedure	*m* (pH 6.5)	Relative yield	Amino acid composition (residues/mol)
T9CR	6.5, 3.5	0.28	++	CmCys (0.4), Thr (0.9), Ser (2.0), Gly (2.1), Ala (1.0), Val (1.4), Phe (1.0)
T9CA1	6.5	0.16	+	Arg (0.8), Asp (1.2), Thr (1.2), Ser (3.0), Glu (2.0), Pro (2.9), Ala (2.7), Val (1.0), Ile (0.9), Leu (1.0), Phe (1.0)
T9CA2	6.5	0.22	++	Ser (3.0), Glu (2.1), Pro (1.9), Val (1.0), Ile (0.8), Leu (0.9), Phe (0.9)
T9CN	6.5, 2.0	0	+	Arg (0.8), Asp (1.0), Thr (0.9), Pro (1.0), Ala (2.9)
T9CB	6.5	–0.28	++	Lys (1.0), Asp (2.0), Pro (1.1), Leu (0.8), Tyr (1.0), Phe (1.0)

peptide Glx-CmCys had a mobility of 1.2, establishing the sequence Asn-Glu-CmCys. Peptide T21 was the only tryptic peptide lacking a basic residue in the *C*-terminal position, and was thus the *C*-terminal peptide.

Chymotryptic peptides of [^{14}C]carboxymethylated light chain

The chymotryptic digest was initially fractionated by paper electrophoresis at pH 6.5. Yields and amino acid compositions (Table 9) were determined from preparative 'fingerprints' as described in the Materials and Methods section. [^{14}C]Carboxymethylcysteine peptides were located by radioautography, whereas other peptides were located by staining with 0.03% (w/v) ninhydrin in acetone. The purification of peptides for sequence determination was as follows. Basic peptides were purified by chromatography in solvent BAWP (for peptides C2, C6 and C7, chromatography was repeated). The neutral band was run at pH 2.0 followed by chromatography in solvent BAWP, and the peptides were eluted after staining with dilute ninhydrin. Acidic bands containing radioactive peptides were oxidized with performic acid and were purified by chromatography in solvent BAWP followed by electrophoresis at pH 3.5, or (for peptides C26 and C29) were purified by electrophoresis at pH 3.5 and oxidation, followed by electrophoresis at pH 3.5. Oxidation converts carboxymethylcysteine into its sulphone derivative, in which the carboxymethyl group has a lower pK_a. The sulphone form is

Table 9. *Amino acid compositions (residues/mol of peptide), yields*

Electrophoretic mobilities are measured relative to aspartic acid at pH6.5, glycylalanine at pH3.5 and valine at pH2.0. Chromatographic R_F in solvent BAWP 0.03% ninhydrin, except for the radioactive carboxymethylated derivatives that were located by radioautography. Purification by paper electrophoresis is indicated

Peptide	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
Purification procedure	6.5, 2.0, BAWP	6.5, BAWP, BAWP	6.5, 2.0, BAWP	6.5, 2.0, BAWP	6.5, BAWP*, 3.5*	6.5, BAWP, BAWP	6.5, BAWP, BAWP	6.5, BAWP	6.5, 2.0, BAWP	6.5, BAWP*, 3.5*	6.5, BAWP	6.5, BAWP*, 3.5*	6.5, 2.0, BAWP	6.5, 2.0, BAWP	6.5, BAWP
Lys		0.9			1.0	1.6	1.5								1.1
His												0.8			
Arg			1.1	0.7				1.0	0.9						
CmCys(O_2)					0.5							0.5			
Asp	1.1		0.2		1.0			1.2	1.0	1.1	2.0				
Thr		0.9	1.0	1.0	1.8				0.8	1.8	0.8			1.0	2.0
Ser		2.2	1.7	0.9	1.1	1.0	1.2	1.1		2.0	2.1		0.9		
Glu		1.1	0.8	1.1	1.2	4.0	4.0				2.1	1.2			
Pro		1.0				2.0	1.9		1.1				1.2		
Gly		0.15	1.2	1.1	0.2			0.7	1.0	2.1		2.2			3.0
Ala					1.0			1.1		1.1	1.9				
Val	0.9		1.9	2.0	1.4				1.0		1.1				
Met	0.9	0.8	0.8												
Ile	0.8										1.0				
Leu			0.9	1.0		2.0	2.0				1.2				1.0
Tyr					0.8	0.7		1.0			0.8	0.9	1.9		
Phe									0.9	1.0				1.0	1.0
Trp															
N-Terminus	Asx	Thr	Ser	Ser	Thr	Tyr	Glx	Gly	Thr	Thr	Thr	His	Ser	Thr	Thr
Mobility (pH6.5)	0.05	−0.31	0	0	0.21	−0.20	−0.22	−0.39	0	0.25	0.54	0.25	0	0	−0.28
Mobility (pH3.5)					0.15*					0.14		0.19*			
Mobility (pH2.0)	0.61		0.81	0.92					0.95				0.63	0.98	
R_F in solvent BAWP	0.96	0.21	0.86	1.15	0.57*	0.54	0.26	0.43	0.88	0.81	1.0	0.41*	1.41	1.32	1.05
% Yield	5.2	4.4	6.3	12.3	24	9.2	12	12	22	30.5	15	20.4	12.5	7.4	9.8

* Indicates a performic acid-oxidized derivative.

thus more acidic at pH3.5, and consequently often easier to separate from other peptides. Peptide C11 was purified by chromatography in solvent BAWP, whereas peptide C23 was purified by electrophoresis at pH3.5. The results of sequence determination by the dansyl–Edman procedure are shown in Fig. 7.

Peptide C1 was the only derivative which, in addition to having an *N*-terminal aspartic acid (or asparagine) residue, was identical with the *N*-terminal sequence of a tryptic peptide (T1). Peptide T1 is therefore the *N*-terminal tryptic peptide.

Peptide C2 overlaps the tryptic peptides T1 and T2, and peptides C3 and C4 overlap the peptides T2 and T3 (Fig. 8). A dipeptide, Ser-Met, was also isolated from the chymotryptic digest but could not be positively placed, as there are three such sequences in the molecule.

Peptide C5 overlaps the tryptic peptides T3 and T4, and peptides C6 and C7 overlap the tryptic peptides T4 and T5 (Fig. 8). The mobilities of these peptides are in agreement with the assignment of amide groups in the tryptic peptide T4. The peptide Val-Ser-Trp was not isolated from the digest, possibly because it had *N*-terminal valine and might therefore have given a poor colour with ninhydrin.

The sequence of peptide C8 was determined as Gly-Ala.... From this and its amino acid composition it must be the *C*-terminal sequence of peptide T5 plus one tyrosine residue. Since peptide T6 is the only tryptic peptide with *N*-terminal tyrosine, peptide C8 must overlap the tryptic peptides T5 and T6. Peptide C9 overlaps the tryptic peptides T6 and T7, as peptide T7 is the only tryptic peptide with *N*-terminal phenylalanine.

Peptides C10, C11, C12 and C13 do not provide any overlaps between different tryptic peptides. However, they provide information about the sequence of peptide T7 (Fig. 5). A dipeptide Thr-Leu was also isolated, but again could not be definitely placed, because there are other similar sequences.

Peptides C15 and C16 showed that peptide T7 was followed by a leucine residue. There are two tryptic peptides T8 and T5, with *N*-terminal leucine. Peptides C6 and C7 placed peptide T5 after peptide T4, and this leaves only peptide T8 as the tryptic peptide after T7. Peptide C17 overlaps the tryptic peptides T8 and T9.

Peptides C18 and C19 do not provide any overlap between tryptic peptides, as they are derived from the middle of peptide T9. The chymotryptic peptide between peptides C17 and C18 was not isolated from the digest of whole light-chain, possibly as it has *N*-terminal valine.

Peptides C20 and C21 provide the overlap between tryptic peptides T10 and T11 (Fig. 8). Peptide T12 was placed after peptide T11, because it is the only tryptic peptide with *N*-terminal tryptophan.

The overlap of peptides T13 and T15 is shown by the amino acid composition and partial sequence of peptide C22 (Figs. 7 and 8). The mobility of peptide

and N-termini of the chymotryptic peptides of MOPC 21 light chain

is relative to valine. Yields are expressed as mol of peptide/mol of protein. All peptides were analysed from chromatographic spots, eluted after reaction with by the pH used, and paper chromatography is indicated by BAWP.

Peptide ...	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29
Purification procedure ...	6.5, BAWP	6.5, 2.0, BAWP	6.5, BAWP*, 3.5*	6.5, 2.0, BAWP	6.5, BAWP	6.5, BAWP	6.5, 2.0, BAWP	6.5, 3.5	6.5, 2.0, BAWP	6.5, BAWP*, 3.5*	6.5, 3.5, 3.5*	6.5, BAWP	6.5, BAWP	6.5, 3.5, 3.5*
Lys	0.8	0.7			1.7	1.8	0.7	0.8		0.8		1.6	1.2	
His											0.8	0.7		
Arg		0.7					0.9							0.9
CmCys(O_2)			0.7								0.6	+		0.6
Asp		1.1		2.0	3.6	2.0	2.1	3.7		1.1				2.0
Thr	1.1	1.0	1.0					2.0	1.0	1.7	1.9	3.4	2.1	
Ser	0.2		2.0				1.0	2.7	2.9			3.0	3.0	
Glu		0.8					2.1	1.2		1.1	1.2	1.2		1.0
Pro		1.0			1.0	0.9						1.1	0.9	
Gly	2.6		2.3				1.9							
Ala		3.2	1.2								1.0	1.2		
Val			1.7		1.2	1.3	0.9					1.2	0.9	
Met									1.0					
Ile		1.0			1.0	0.8	1.0					0.8	0.8	
Leu	1.0			1.0	1.1		1.0		1.0	0.9				
Tyr					0.9	0.9		0.7		0.9				
Phe			0.9	0.9	1.0							0.9	1.0	
Trp					+	+		+						
N-Terminus	Gly	Glx	Thr	Leu	Leu	Tyr	Lys	Asx	Ser	Thr	Thr	Thr	Lys	Asx
Mobility (pH6.5)	−0.42	0	0.26	0	−0.18	−0.24	0	0.40	0	0.25	0.42	−0.1	−0.45	0.40
Mobility (pH3.5)			−0.42*					0.28		0.89	0.67			1.15
Mobility (pH2.0)		1.05		0.65			1.01		0.63					
R_F in solvent BAWP	0.54	0.25	0.92*	1.10	0.97	0.83	0.55		0.82	0.88		0.23	0.61	
% Yield	11.5	17.3	22.1	5.0	3.7	7.2	4.8	4.2	17.4	8.8	17.5	3.2	11.2	18.5

C22 (neutral) together with the mobility of peptide T13 (acidic), permits the assignment of two amide groups in the sequence Gln-Asn-Gly-Val ...

Peptide C23 provides an overlap between peptides T15 and T16. The sequence of peptides C24 and C25 was used to complete the sequence of peptide T16. In addition, peptide C25 overlaps the peptides T16 and T17.

Peptides C26, C27 and C28 gave evidence for overlaps between the tryptic peptides T18, T19 and T20 (Fig. 8). The electrophoretic mobility of peptide C26 indicates the presence of more than one net acidic charge, indicating a glutamic acid residue rather than the amide; the mobility of peptide C27 indicated a fractional basic charge, confirming the assignment.

Peptide C29 provided the overlap between tryptic peptide T20 and the *C*-terminal peptide T21.

The chymotryptic peptides therefore provided a unique way in which all the tryptic peptides could be arranged into three segments, which included the tryptic peptides T1 and T12, T13 to T17, and T18 to T21. Since peptide T1 is the *N*-terminal and peptide T21 the *C*-terminal end of the molecule, peptides T12 to T17 must be between the other two segments, as shown in Fig. 8. There is in fact one residue (lysine) in peptide C22, which supports this scheme and overlaps peptides T12 and T13. Unfortunately, a chymotryptic peptide overlapping peptides T17 and T18 was not detected.

The amino acid composition determined for the light-chain is in very good agreement with that calculated from the sequence (Table 1). Thus, it is unlikely that there are any undetected tryptic peptides. Tryptophan was not determined in isolated peptides, but the number deduced from the sequence agrees well with the tryptophan content determined for the whole protein.

Discussion

Technical aspects of the sequence determination

The derivation of the sequence of MOPC 21 light-chain was facilitated by the knowledge of the sequences of light chains from other sources. Previous studies on other mouse κ-chains (Gray *et al.*, 1967) were useful, especially in the elucidation of the *C*-terminal half of the chain. However, as those studies were incomplete it was considered important to complete the sequence as far as possible.

The availability of highly sensitive amino acid analysers and even more sensitive degradation procedures makes it seem paradoxical that, at present, the amount of peptide required for sequence determination is often of the same order of magnitude as the amount usually discarded during ninhydrin-staining procedures. In this study, we have often eluted peptides after a dilute ninhydrin stain (Anfinsen *et al.*, 1959). The peptides usually show a 20–30% loss of the

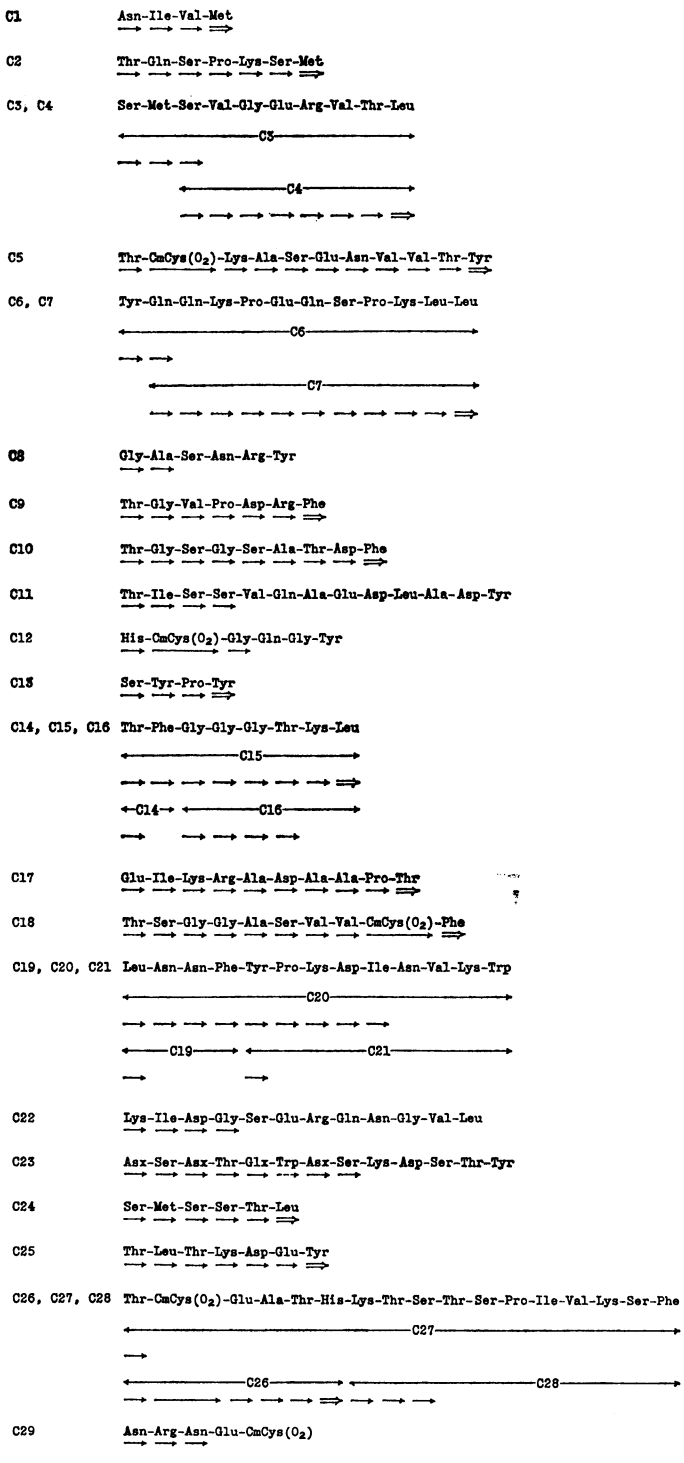

Fig. 7. *Sequence determination of chymotryptic peptides*

Amino acid compositions are shown in Table 9. For details see the text.

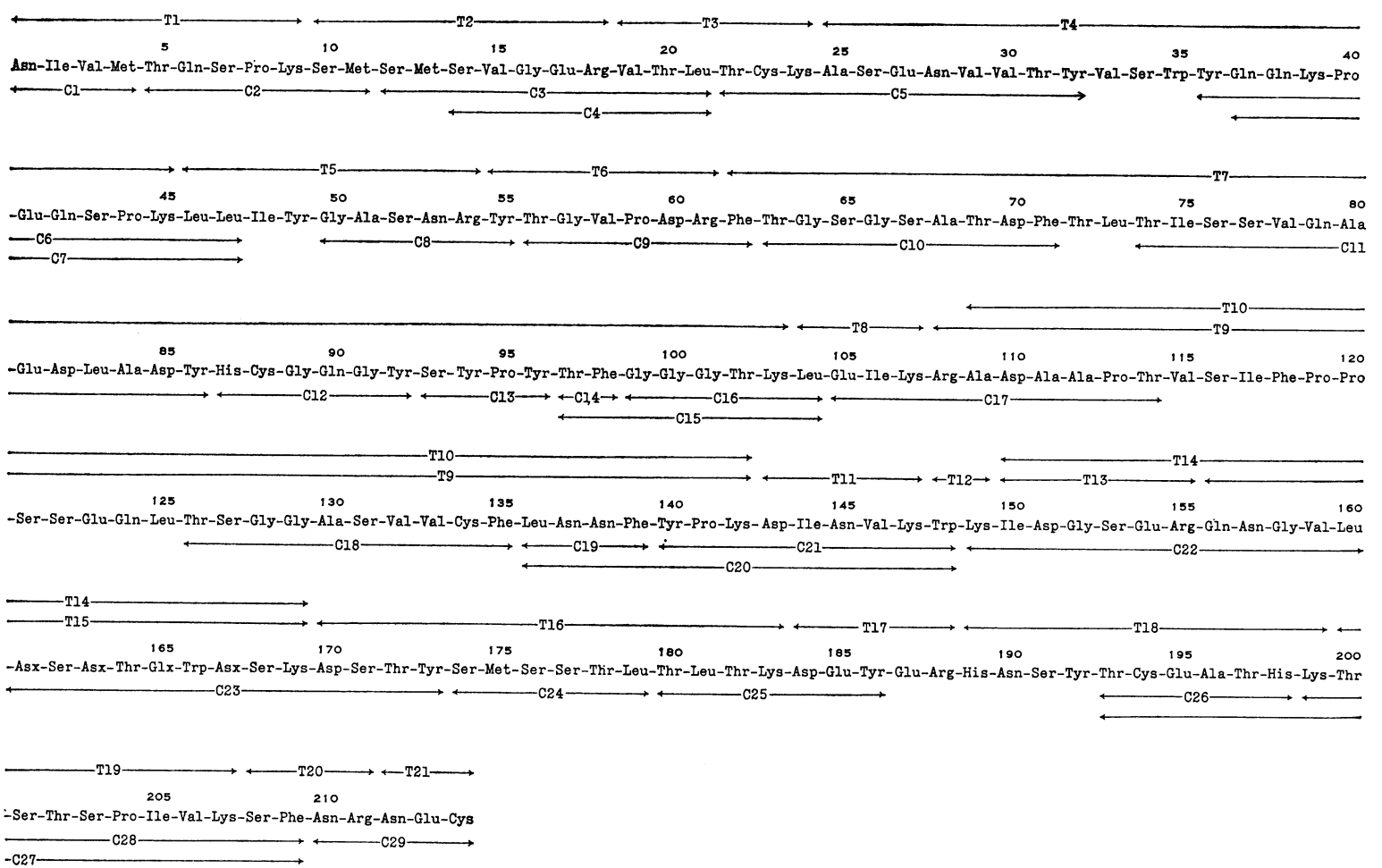

Fig. 8. *Amino acid sequence of MOPC* 21 *light-chain, showing the tryptic peptides (names beginning with T) and chymotryptic peptides (names beginning with C)*

Cystine peptides were isolated as [^{14}C]carboxymethylcysteine (sulphone) derivatives. Cys-214 has been shown to be disulphide-bonded to heavy chain; in addition, disulphide bridges join Cys-23 to Cys-88, and Cys-134 to Cys-194 (Svasti & Milstein, 1972). The assignment of amide residues in peptides T14, T15 and C23 is extensively discussed in the text.

N-terminal residue and of lysine. The eluted material could be subjected to the usual dansyl–Edman degradation without more ambiguities or difficulties than we would have expected with more conventional handling procedures. The cases where this has been done are mentioned in the text.

Two peptides (T1 and T15) were isolated with different mobilities at pH 6.5 (see 'Tryptic peptides' in the Results section). Deamidation of the *N*-terminal asparagine residue seems to be the obvious explanation for the charge differences in peptide T1. In peptide T15, cyclization of the *N*-terminal glutamine into pyrrolid-2-one-5-carboxylic acid, in addition to deamidation, is likely to have contributed to the multiple forms obtained. Of the two peptides, T1 was the more sensitive to deamidation. The possibility that deamidated residues in peptides T1 and T15 are already present in serum cannot be discarded. Microheterogeneity in myeloma proteins is always observed in serum and this has been ascribed to fast deamidation by exposure to serum (Awdeh *et al.*, 1970).

With the above proviso, MOPC 21 light chain appears as a homogeneous single sequence of 214 residues. A disulphide-bridged peptide has been obtained from a peptic–tryptic digest of whole myeloma protein (Svasti & Milstein, 1972), which demonstrates a bond between half-cystine residues at positions 134 and 194. In the same paper it was shown that cysteine-214 was involved in interchain binding, and as the protein did not have free thiol groups, the remaining two half-cystine residues in positions 23 and 88 must form a second intrachain bond.

Significance of the sequence of MOPC 21 light chain in relation to other immunoglobulin chains

It is now well recognized that the light chains of human myeloma proteins contain a section (the *C*-terminal half), which is an essentially invariant feature of proteins of the same type, whereas considerable variation occurs in the *N*-terminal half (reviewed in Cohen & Milstein, 1967; Lennox & Cohn, 1967). 'Fingerprint' studies (Bennett *et al.*, 1965) and partial sequence studies (Gray *et al.*, 1967) indicated that this was also the case in mouse κ-chains. The present studies confirm this conclusion. There are, however,

Table 10. *Differences (minimum number of base changes/100 residues compared) between human and mouse C_L-regions, human κ basic sequences and $V_κ$-regions of MOPC 21, MOPC 41 and MOPC 70 proteins*

A gap is calculated as two base changes. References: basic sequences are as described by Milstein & Munro (1970), complemented by the results of Milstein & Deverson (1971); human C_L sequences are taken from Milstein & Pink (1970); the mouse $C_λ$ sequence is from Appella (1971); the V-region sequences are of MOPC 41 and MOPC 70 from Gray *et al.* (1967); the mouse $C_κ$ and V-region sequence of MOPC 21 are from the present paper.

(*a*) Human and mouse C_L-region differences

	Human $C_κ$	Mouse $C_κ$	Human $C_λ$
Mouse $C_κ$	52		
Human $C_λ$	87	82	
Mouse $C_λ$	90	90	37

(*b*) Human κ basic sequences and $V_κ$-regions of mouse proteins

	MOPC 21	MOPC 41	MOPC 70	Human$_{κIa-b}$	Human$_{κII}$
MOPC 41	55				
MOPC 70	53	54			
Human $V_{κIa-b}$	42–43	37–35	45–44		
Human $V_{κII}$	45	53	46	41–46	
Human $V_{κIII}$	46	47	44	37–40	32

a few differences between the *C*-terminal half of the two previous κ-chain sequences (Gray *et al.*, 1967) and the one given in the present paper. Those which occur at residues 129, 161 and 165 are not real differences, as the sequences in those sections were deduced by Gray *et al.* (1967) by homology with human chains and were within parentheses. The only real difference in sequence occurs at residue 127.

Differences in the *C*-terminal half have been shown to be either due to genetic polymorphism (e.g. human κ-chains, reviewed by Milstein & Pink, 1970), to gene multiplicity (e.g. human λ-chains, Ein, 1968) or to technical problems associated with sequence determination. The first possibility is most unlikely, as these tumours are derived from the same inbred strain of mice. The second and third possibilities cannot be excluded without further investigation.

A revised comparison of the *C*-regions of mouse κ-, mouse λ-, human κ- and human λ-chains (Table 10a) confirms the generally held view that κ- and λ-chain genes have diverged early in evolutionary history. The difference in sequence between human and mouse κ-chains is now 52%, which is considerably higher than the difference between human and mouse λ-chains (37%). This indicates that the rate of evolution of κ-chains is significantly higher than that of λ-chains.

It has been suggested that the pattern of variability of the sequence in the *N*-terminal half of human κ-chains is the result of non-random somatic mutation events on a restricted number of genes present in the germ line. Such a process would explain the small number of well-defined basic sequences found, for instance, in human κ-chains (Milstein, 1967; reviewed in Milstein & Munro, 1970). However one basic sequence may be expressed by more than a single gene (Milstein *et al.*, 1969; Gibson *et al.*, 1971). This fact, added to the postulated 'species specificity' of basic sequences has been explained by an evolutionary process of expansion and contraction (Milstein & Pink, 1970). According to this suggestion, the number of genes coding for V-regions is likely to vary in different species and even, to some degree, within species. Studies on the variability of the *N*-terminal half of mouse κ-chains are therefore of interest both in deriving the basic sequences presumed to be present in mice and in correlating these basic sequences with those of human κ-chains.

Table 10(b) shows a comparison of the kappa V-regions of MOPC 21, MOPC 41 and MOPC 70 proteins with the basic sequences of the human V_κ regions. In evaluating these results, it must be remembered that comparisons between mouse proteins include differences in the 'hot spots' (i.e. regions of high variability) and less conservative positions, whereas comparisons involving basic sequences do not. The results nevertheless suggest very strongly that the three mouse sequences belong to different (as yet undefined) basic sequences. In addition, MOPC 21 and MOPC 70 proteins do not seem to be more similar to any one of the human basic sequences, and differ about as much from other mouse proteins as they do from the human basic sequences. This suggests that the ancestral V-genes, which gave rise to these two proteins, already existed before the separation of the mouse and human evolutionary lines.

Based on studies of the *N*-terminal sequences of 22 Balb/c κ-chains, Hood *et al.* (1970) have suggested a minimum of nine subgroups. The equivalent sequence of MOPC 21 immunoglobulin is so different from the nine subgroups, that it should belong to a new subgroup. It must be emphasized, however (Milstein & Pink, 1970; Milstein & Svasti, 1971), that the definition of subgroups based on individual proteins cannot be more than tentative until more extensive results are available. If out of 23 proteins tested ten subgroups are detected and some of these subgroups are represented by single proteins, it is statistically likely that the number of subgroups not yet detected is considerable. This number is difficult to calculate because, as found in humans (Milstein *et al.*, 1969), the contribution of each subgroup in the total population is likely to vary considerably. Therefore restricted repeats are expected to be frequent if a few subgroups make up a good proportion of the chains, and the rest is made up by a number of other subgroups each representing a small fraction of the total population.

We thank Mr. F. Northrop for technical assistance. J. S. gratefully acknowledges the receipt of a Coutts-Trotter research fellowship from Trinity College, Cambridge.

References

Ambler, R. P. (1963) *Biochem. J.* **89**, 341–349
Ambler, R. P. & Brown, L. H. (1967) *Biochem. J.* **104**, 784–825
Anfinsen, C. B., Åqvist, S. E. G., Cooke, J. P. & Jönssön, B. (1959) *J. Biol. Chem.* **234**, 1118–1123
Appella, E. (1971) *Proc. Nat. Acad. Sci. U.S.* **68**, 590–594
Awdeh, Z. L., Williamson, A. R. & Askonas, B. A. (1970) *Biochem. J.* **116**, 241–248
Bennett, J. C., Hood, L., Dreyer, W. J. & Potter, M. (1965) *J. Mol. Biol.* **12**, 81–87
Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228
Cohen, S. & Milstein, C. (1967) *Advan. Immunol.* **7**, 1–89
Crowshaw, K., Jessup, S. J. & Ramwell, R. W. (1967) *Biochem. J.* **103**, 79–85
Delange, R. J. & Smith, E. L. (1971) *Enzymes*, 3rd. edn., **3**, 81–118
Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–95
Ein, D. (1968) *Proc. Nat. Acad. Sci. U.S.* **60**, 982–985
Gibson, D., Levanon, M. & Smithies, O. (1971) *Biochemistry* **10**, 3114–3122

Gray, W. R. (1967*a*) *Methods Enzymol.* **11**, 139–151
Gray, W. R. (1967*b*) *Methods Enzymol.* **11**, 465–375
Gray, W. R., Dreyer, W. J. & Hood, L. (1967) *Science* **155**, 465–467
Heilmann, J., Barrolier, J. & Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219
Hirs, C. H. W. (1967) *Methods Enzymol.* **11**, 197–199
Hood, L. E., Potter, M. & McKean, D. J. (1970) *Science* **170**, 1207–1210
Lennox, E. S. & Cohn, M. (1967) *Annu. Rev. Biochem.* **36**, 365–406
Milstein, C. (1966) *Biochem. J.* **101**, 338–351
Milstein, C. (1967) *Nature* (*London*) **216**, 330–332
Milstein, C. (1968) *Biochem. J.* **110**, 652–654
Milstein, C. & Munro, A. J. (1970) *Annu. Rev. Microbiol.* **24**, 335–358
Milstein, C. & Pink, J. R. L. (1970) *Progr. Biophys. Mol. Biol.* **21**, 211–263
Milstein, C. & Svasti, J. (1971) *Progr. Immunol.* (Amos, B., ed.), pp. 33–45, Academic Press, New York and London
Milstein, C., Milstein, C. P. & Franklin, E. C. (1969) *Nature* (*London*) **221**, 151–154
Milstein, C. P. & Deverson, E. V. (1971) *Biochem. J.* **123**, 945–958
Offord, R. E. (1966) *Nature* (*London*) **211**, 591–593
Pink, J. R. L., Buttery, S. H., de Vries, G. M. & Milstein, C. (1970) *Biochem. J.* **117**, 33–47
Plummer, T. H. (1969) *J. Biol. Chem.* **244**, 5246–5253
Sanger, F. & Thompson, E. O. P. (1963) *Biochim. Biophys. Acta* **71**, 468–471
Schiechl, H. & Hilschmann, N. (1971) *Hoppe-Seyler's Z. Physiol. Chem.* **352**, 111–115
Smith, G. P., Hood, L. & Fitch, W. M. (1971) *Annu. Rev. Biochem.* **40**, 969–1012
Smith, I. (1953) *Nature* (*London*) **171**, 43–44
Spies, J. R. & Chambers, D. C. (1948) *Anal. Chem.* **20**, 30
Svasti, J. & Milstein, C. (1972) *Biochem. J.* **126**, 837–850
Waley, S. & Watson, J. (1953) *Biochem. J.* **55**, 328–337
Wang, A. C., Pink, J. R. L., Fudenberg, H. H. & Ohms, J. J. (1970) *Proc. Nat. Acad. Sci. U.S.* **66**, 657–663
Weigert, M. G., Cesari, I. M., Yonkovitch, S. J. & Cohn, M. (1970) *Nature* (*London*) **228**, 1045–1047
Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta* **133**, 369–370