AMINO ACID SEQUENCE OF A LAMBDA BENGE JONES PROTEIN FROM A CASE OF PRIMARY AMYLOIDOSIS

NOBUHIRO TAKAHASHI, TATSUNORI TAKAYASU, TOMOTAKA SHINODA¹, SHIGENORI ITO², TSUNEO OKUYAMA and AKIRA SHIMIZU

Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Setagaya-ku, Tokyo 158 and Department of Internal Medicine, Osaka University School of Medicine, Fukushima-ku, Osaka 553, Japan

ABSTRACT

Amino acid sequence analysis was done on a λ Bence Jones protein NIG-51 from a case of primary amyloidosis with the major objective of determining the sequence of the variable half of the molecule. Nineteen major tryptic peptides covering 216 residues were isolated. The complete sequences of 15 peptides of these, comprising the entire variable half and 7 from the constant half, were determined. For the remaining peptides covering the rest of the C-terminus, only partial sequences or the amino acid compositions were determined. On the basis of the data all the tryptic peptides could be arranged in order. The deduced sequence of the C-terminal half corresponds exactly to that established for human λ light chains of the same isotype. The V region includes two extra residues (27b-27c) and is characteristic of the V λ I subgroup.

A brief discussion of the possible relationship between these replacements and the formation of tissue deposits, and subgroup specificity is also presented.

KEY WORDS amino acid sequence / Bence Jones protein / primary amyloidosis / V λ I subgroup

Primary amyloidosis is a disease characterized by extracellular deposits of a microscopically homogeneous fibrillar AL protein in various tissues (2). Occasionally such patients excrete into the urine a protein characteristic of Bence Jones protein or its fragments as in the case of multiple myeloma (7, 15). Although a large body of sequence data on Bence Jones proteins as well as on light chains has been reported to date, only limited and incomplete sequence data are available on AL proteins, which seem to be insufficient to elucidate the relationship between primary structure and pathogenic functions of such proteins.

This article presents the complete sequence of the variable region of the Bence Jones protein NIG-51 of subgroup V λ I from a case of primary

amyloidosis.

MATERIALS AND METHODS

Materials

Bence Jones amyloid protein NIG-51 was isolated from the urine of a patient with primary amyloidosis. TPCK-treated trypsin (EC 3.4.21.4) and chymotrypsin (EC 3.4.21.1) were obtained from Whorthington Biochemical Co. Acid carboxypeptidase from *penicillium janthinellum* (CPase P; EC 3.4.23.7) was obtained from Takara Shuzo (Shiga, Japan). Specific antisera were purchased from Boehringerwerke A.G. DEAE-Sephadex A-50 (40-120 μ), A-25 (40-120 μ) and Sephadex G-25 (10-40 μ) were purchased from Pharmacia. 3N Mercaptoethane sulfonic acid was purchased from Pierce Chemical Co. Reagents and solvents for Edman degradation (sequence grade) and other reagents (analytical grade) were obtained from Wako

1. To whom correspondence should be addressed

2. Present address: Department of Legal Medicine, Osaka Medical University, Daigakucho, Takatsuki 569, Japan

Pure Chemical Industries.

Purification of Protein NIG-51

The urine was concentrated by lyophilization and dialyzed against 0.01 M phosphate buffer containing 0.05% NaN₃, pH 8.0. The dialyzed sample was applied to a column of DEAE-Sephadex A-50 (3.6 × 5 cm) equilibrated with 0.01 M phosphate buffer containing 0.05% NaN₃, pH 8.0 and was eluted with the equilibration buffer. Homogeneity was tested by disc-electrophoresis, SDS-electrophoresis, immunoelectrophoresis, immunodiffusion and N-terminal analysis.

Polyacrylamide Gel Electrophoresis

Disc-electrophoresis was performed in a glass tube with 7.5% gel as described by Ornstein (11) and Davis (3). The gel was stained with 0.025% Coomassie Brilliant Blue G-250 in 50% (v/v) methanol-7% (v/v) acetic acid solution overnight and destained by several changes of 7% acetic acid for 2 days. SDS-disc-electrophoresis was performed according to Dunker and Ruckert (4). The gel was stained with 0.01% Coomassie blue G-250-50% methanol-7% acetic acid (v/v).

Immuno-electrophoresis and Immunodiffusion

Immuno-electrophoresis and immunodiffusion were carried out by the method of Ouchterlony (12), using an agar gel plate of 2.5 × 7.5 cm. Immuno-electrophoresis was carried out at a constant current of 2 mA/cm.

Reduction and Aminoethylation

Purified protein (235 mg) was dissolved in 10 ml of 1M Tris-HCl (pH 8.6) containing 8M urea and 2 mM EDTA, and then 0.4 ml of 2-mercaptoethanol was added. The mixture was kept under nitrogen for 4 hr at room temperature to allow complete reduction. Then 0.6 ml of ethylenimine was added, and the mixture was kept at room temperature for 1 hr to allow aminoethylation. The resulting mixture was dialyzed against 0.1 M NH₄HCO₃ at 4°C.

Tryptic Digestion and Fractionation of the Digest

The aminoethylated protein (100 mg) in 0.1 M NH₄HCO₃ was digested with 1.5 mg of TPCK-

treated trypsin for 2 hr at 37°C. The digest was lyophilized and suspended in 5 ml of 0.01 M NH₄HCO₃. A minute amount of insoluble material was removed by centrifugation at 28,000 *g* for 10 min. The precipitate was washed twice with 2.5 ml of 0.01 M NH₄HCO₃. The supernatants were combined and applied to a column of DEAE-Sephadex A-25 (1.5 × 42 cm) equilibrated with 0.01 M NH₄HCO₃. The column was eluted at a flow rate of 20 ml/hr, first with 360 ml of 0.01 M NH₄HCO₃ and then with a linear gradient (800 ml each) of increasing ammonium bicarbonate concentration (0.01 M–0.6 M), followed by further elution with a linear gradient (300 ml each) of increasing ammonium bicarbonate concentration (0.6 M–1.0 M). Fractions of 5.3 ml were collected. Homogeneity of the peptides was tested by amino acid analysis after hydrolysis, and Edman degradation. Further purification of peptides was performed by gel filtration on a column of Sephadex G-25 using 0.1 M NH₄HCO₃ buffer at room temperature.

Amino Acid Analysis

Protein and peptides were hydrolyzed with 6N HCl in evacuated, sealed tubes at 110°C for 24 hr. Analysis was performed with an automatic amino acid analyzer (Hitachi KLA-5) using the method of Spackman *et al.* (21).

Tryptophan contents in the protein and peptides were determined by amino acid analysis after hydrolysis with 3N mercaptoethane sulfonic acid, as described by Penke *et al.* (13). No corrections were made for destruction of serine, threonine, and tyrosine in the protein and peptides during acid hydrolysis.

Sequence Analysis

The manual Edman degradation was carried out by the conventional procedure with minor modifications (8). PTH amino acids were identified by liquid chromatography according to the procedure of Isobe and Okuyama (8). The PTH derivative of histidine was also identified by liquid chromatography, in a column of Partisil 10 (4 mm × 15 cm) at a flow rate of 1.3 ml/min. Elution was carried out with MeOH:AcOH:H₂O (75:5:20) (v/v). PTH histidine was eluted at 4.42 min.

Digestion with CPase P was performed in 0.2 M formic acid (adjusted to pH 2.5 with pyridine) (29). Aliquots of the digestion mixture were taken at intervals, and subjected to amino acid

analysis.

Tryptic peptide TV1 (1.4 mg) with a blocked N-terminus was digested with 60 μ g of chymotrypsin in 0.2 ml of 0.1 M NH_4HCO_3 , pH 8 for 2 hr at 37°C for further sequence analysis. The digest was chromatographed on a column of Sephadex G-25 (1.5 $\times$ 90 cm) in 0.1 M NH_4HCO_3 , pH 8 at a flow rate of 5.7 ml/hr.

Nomenclature of Peptides

The nomenclature of peptides of the protein NIG-51 follows that adopted for the other λ chains (16, 24, 27). Tryptic peptides are given the prefix T and are identified by the letter V for the variable region and λ for the constant region of the sequence. Since the protein was aminoethylated, additional tryptic peptides ending with AE-cysteine were obtained; these are given the same number as for the tryptic peptide expected from the unsubstituted protein plus the letter *a*, and the remaining tryptic peptide has the same number plus the letter *b*. The peptides in the variable region in the protein are numbered consecutively in their order in the sequence. Tryptic peptides in the constant region are numbered consecutively beginning with hypothetical T λ 1.

RESULTS

Purity and Amino Acid Composition of the Protein NIG-51

The purified protein gave a single band both on disc-electrophoresis and SDS-disc-electrophoresis. The protein also showed a single arc on immunoelectrophoresis against anti-human serum. On immunodiffusion, it gave a positive reaction to only anti- λ chain antiserum and a negative reaction to other antisera, such as anti- κ chain, anti-transferrin and anti-albumin.

The N-terminus of the protein was not detected by Edman degradation, indicating that it was blocked. No significant amounts of other N-terminal amino acids were detected. Therefore, the final preparation was considered pure enough for sequence analysis.

The amino acid composition of the aminoethylated protein is summarized in Table 1.

Purification of Tryptic Peptides

The tryptic peptides were separated into sixteen fractions on a DEAE-Sephadex A-25 column, as shown in Fig. 1. The fractions 1, 4, 5, 6, 11, 12, 13, 14, 15 and 16 were already pure and each

fraction contained TV1b, TV2, T λ 3a, T λ 5, T λ 3b, TV3, T λ 7, TV5b, TV5a and TV5ab, respectively. The other fractions were further purified by gel filtration on a Sephadex G-25 column, as shown in Fig. 2, A–F. Fraction 2, Fig. 1 contained TV1a, TV6-T λ 1 and T λ 8 (Fig. 2, A). Fraction 3 contained TV4, TV6-T λ 1 and T λ 6 (Fig. 2, B). Fraction 7 contained T λ 4 and T λ 2 (Fig. 2, C). Fraction 8 contained TV3 and T λ 2 (Fig. 2, D). Fraction 9 contained T λ 9a and T λ 10 (Fig. 2, E). Fraction 10 contained T λ 9b (Fig. 2, F).

In all, twenty tryptic peptides were obtained from the tryptic digest of the aminoethylated protein. The amino acid compositions, number of residues and the yields of the twenty peptides are summarized in Table 1.

An additional peptide, TV5ab was obtained from the digest due to incomplete cleavage by trypsin of a bond containing an aminoethyl-cysteine residue (17). This peptide gave overlap for two tryptic peptides, TV5a and TV5b. Summation of residues found in the peptides accounts for the amino acid composition of the aminoethylated protein.

Amino Acid Sequence of Tryptic Peptides

In this paper the alignment of the peptides is presented in the way adopted by Titani *et al.* (24) and Shinoda *et al.* (16). The proof of the sequence or partial sequence is given in the order that the peptides appear in the primary structure of the protein. The alignment of tryptic peptides was deduced in part by sequencing overlapping tryptic peptides as well as by homology with the corresponding regions of other λ chains of the same subgroup. Altogether, the complete sequence of 15 peptides, 8 from the variable region and 7 from the constant region, were determined. The complete sequence of the remaining peptides from the constant region was not determined, because all of these were readily identifiable by their homologies with those of the constant region of protein Sh (24, 25, 28) on the basis of various properties, such as amino acid composition, N-terminal, and partial sequences.

TV1a Among twenty tryptic peptides, three peptides, TV1a, T λ 6 and T λ 9b, did not undergo Edman degradation as in the original protein. While two of these peptides, T λ 6 and T λ 9b, were identified as homologous with those of the constant region of protein Sh (25), the remaining peptide, TV1a, did not accord with any of the tryptic peptides of the constant region of protein Sh on the basis of amino acid composition shown

Table 1 *Amino Acid Composition of Tryptic Peptides of the Aminoethylated Protein NIG-51*

	TV1a	TV1b	TV2	TV3	TV4	TV5a	TV5b	TV5a-b	TV6-Tλ1	Tλ2	Tλ3a
Lys			1.0		1.2		1.0	1.0	0.8	1.0	
His						1.0		0.9			
Arg		1.0		0.8							
AECys ^a	0.6					0.9		0.8			0.7
Asp		1.3	1.8	2.8		2.4	3.4	5.1		1.1	
Thr	1.0		1.1			1.1	1.7	2.9	1.0	1.3	1.0
Ser	4.8	4.2		1.8	2.0	4.9	1.4	5.4		2.9	
Glu	3.3		2.3	1.1		2.3		2.2	1.1	3.3	
Pro	3.2		2.4	2.0			1.1	1.1	1.1	3.0	
Gly	2.0	2.0	1.2	1.0	1.0	2.3	4.0	6.1	1.0		
Ala	0.9		2.0			3.1	0.9	4.0		3.0	1.0
Val	2.4		2.0	2.0			1.3	1.0	1.0	1.2	0.9
Met											
Ile	1.4	1.0				1.0		1.0			
Leu	1.0			1.7		2.1	1.4	3.0	1.9	2.0	0.9
Tyr			1.0	0.8		0.8		0.9			
Phe					1.0	1.1	0.8	1.9		1.0	
Trp			0.7 ^b	0.8 ^b			0.7 ^a	0.7 ^b			
Total residues	22	9	15	16	5	22	17	39	8	19	5
Yield %	45	91	68	26	79	10	11	53	40	70	46
	Tλ3b	Tλ4	Tλ5	Tλ6	Tλ7	Tλ8	Tλ9a	Tλ9b	Tλ10	NIG-51 ^c	
Lys	1.0	1.0	1.1	1.0	0.9			0.9		10.7(11)	
His						0.9		0.8		2.7(3)	
Arg						1.0				2.8(3)	
AECys ^a							0.9		0.6	4.5(5)	
Asp	1.1	1.0		1.8						15.4(16)	
Thr	0.9		2.8		1.1			1.7	2.0	16.0(17)	
Ser	1.1	1.4	0.9	0.9	2.9	0.8	2.0	1.2	0.9	30.5(35)	
Glu			1.1	1.0	2.2			2.9	1.4	19.9(20)	
Pro	1.1	1.1	1.0		1.1				1.4	18.4(17)	
Gly	1.0		1.0					1.0		17.0(17)	
Ala	1.9	0.8	1.0		2.0				1.0	18.0(18)	
Val	2.1	1.0	1.1					2.1	1.3	17.3(18)	
Met										0 (0)	
Ile	0.9									4.3(5)	
Leu	1.0				2.0					14.1(14)	
Tyr	0.9				2.1		1.0			6.9(7)	
Phe	1.0									4.8(5)	
Trp	0.9 ^b				0.9 ^b					5.1 ^b (5)	
Total residues	15	7	10	5	15	3	4	11	8	(216)	
Yield %	46	48	52	37	87	41	39	26	77		

a. AECys refers to aminoethyl cysteine.

b. Tryptophan was determined by amino acid analysis after hydrolysis with 3N mercaptoethane sulfonic acid.

c. Amino acid composition of aminoethylated protein NIG-51. The numbers of residues found by sequence analysis are shown in parentheses.

Fig. 1 Purification of the tryptic peptides of the aminoethylated protein on a column (1.5 × 42 cm) of DEAE-Sephadex A-25. The tryptic peptides from 100 mg of the protein, dissolved in 0.01 M NH_4HCO_3 , were applied to the column equilibrated with 0.01 M NH_4HCO_3 . The column was eluted at a flow rate of 20 ml/hr, first with 360 ml of 0.01 M NH_4HCO_3 and with a linear gradient (800 ml each) of increasing ammonium bicarbonate concentration (0.01 M–0.6 M), followed by further elution with a linear gradient (300 ml each) of increasing ammonium bicarbonate concentration (0.6 M–1.0 M). Each 5.3 ml of the eluate was collected. Peptides were pooled as indicated by horizontal bars.

Fig. 2 Purification of fractions from Fig. 1 on a column (1.5×90 cm) of Sephadex G-25. The fractions from Fig. 1, dissolved in 0.5 ml of ammonia water, were applied to the column equilibrated with 0.1 M NH_4HCO_3 . The column was eluted with 0.1 M NH_4HCO_3 at a flow rate of 5.7 ml/hr, and 2.2 ml-fractions were collected. Peptides were pooled as indicated by horizontal bars. (A) fraction 2 from Fig. 1; (B) fraction 3; (C) fraction 7; (D) fraction 8; (E) fraction 9; (F) fraction 10

in Table 1. Therefore, TV1a was deduced to be the N-terminal tryptic peptide of the protein.

For further sequence analysis, the peptide TV1a was digested with chymotrypsin and the

digest was separated into two fractions on a Sephadex G-25 column (Fig. 3). Fractions 1 and 2 contained TV1aC2 and TV1aC1, respectively. The amino acid compositions of the

chymotryptic peptides are shown in Table 2. The combined composition accounts for the amino acid composition of TV1a. The alignment of the two peptides was determined to be (TV1aC1)-(TV1aC2) because the N-terminus of TV1aC1 was not detected by Edman degradation, while that of TV1aC2 was threonine and its C-terminal sequence was Ile-Ser-AEC as in the original peptide. TV1aC1 contained four amino acid residues and CPase P digestion of the peptide gave -Val-Leu as the C-terminal sequence. The N-terminus of TV1aC1, like TV1a, was deduced to be pyrrolidone carboxylic acid on the basis of homology with the corresponding portion of other λ chains (16, 27). TV1aC2 comprising eighteen residues was sequenced throughout by Edman degradation. The AE-cysteine present in the peptide TV1a is the first half-cystine in the variable region of the protein NIG-51.

TV1b This peptide contains nine residues, and was sequenced throughout by Edman degradation. The mean repetitive yield of eight cycles of the degradation was on the average 89%. The remaining residue after eight cycles of the degradation was applied to an amino acid analyzer without hydrolysis and free arginine

was detected. The sequence was homologous with the region from the 22nd to 28th residues in the protein Ha (16), and thus it was deduced that the peptide occupied the corresponding region of the protein NIG-51.

TV2 This peptide containing a tryptophan residue consists of fifteen residues. The sequence of the first fourteen residues was determined by Edman degradation. The C-terminal residue was deduced to be lysine because of the specificity of trypsin. The determined sequence was homologous with that of the last fifteen residues of TVh2b of protein Ha (16), and thus it was deduced that the peptide occupied the corresponding region of the protein NIG-51.

TV3 This peptide containing a tryptophan residue consists of sixteen residues, and was sequenced throughout by Edman degradation. The repetitive yield of fifteen cycles of the degradation was on the average 85%. The remaining residue after fifteen cycles was applied to an amino acid analyzer without hydrolysis, and free arginine was detected. The sequence was homologous with that of TVh3-TVh4 of protein Ha (16), and thus it was deduced that the peptide occupied the corresponding region of the protein NIG-51.

TV4 This peptide contains five residues. Four cycles of Edman degradation and CPase P digestion provided the sequence shown in Table 3. The sequence was identical with that of TVh5 of protein Ha (16).

TV5a This peptide contains twenty-two residues. Twenty-one cycles of Edman degradation and CPase P digestion provided the sequence shown in Table 3. The sequence was identical with that of TVh6-TVh7a from protein Ha (16), except that Arg-77, His-83, and His-85 of the latter were respectively replaced by His, Asp, and Phe in the present sample, and thus it was deduced that the peptide occupied the corresponding region of the protein NIG-51.

TV5b This peptide containing a tryptophan residue consists of seventeen residues. Sixteen cycles of Edman degradation and CPase P digestion provided the sequence shown in Table 3. The sequence was homologous with the region containing TVh7b and the first eleven residues of TVh8 of protein Ha, and thus it was deduced that the peptide occupied the corresponding region of the protein NIG-51.

TV5a-b This peptide containing AE-cysteine and lysine consists of thirty-nine residues. The sequence of the first twenty-seven residues was determined by Edman degradation. Although the PTH derivative of step twenty-two of the

Table 2 Amino Acid Composition of Chymotryptic Peptides from Tryptic Peptide TV1a

	TV1aC1	TV1aC2	TV1a
Lys			
His			
Arg			
AECys*		0.7	0.6
Asp			
Thr		1.0	1.0
Ser	1.1	3.7	4.8
Glu	1.2	2.2	3.3
Pro		3.5	3.2
Gly		2.1	2.0
Ala		1.0	0.9
Val	1.0	1.6	2.4
Met			
Ile		1.4	1.4
Leu	1.0		1.0
Tyr			
Phe			
Trp			
Total residues	4	18	22
Yield %	84	76	

* Aminoethyl cysteine

degradation was eluted at a position identical to PTH lysine in liquid chromatography, it was deduced to be PTH AE-cysteine because the lysine residue in the peptide occupied the C-terminal position. CPase P digestion gave lysine as its C-terminus. The amino acid composition of TV5a-b was in complete agreement with the summation of residues found in the sequences of TV5a and TV5b. Therefore, the peptide contained two tryptic peptides, TV5a and TV5b, with alignment (TV5a)-(TV5b), thus giving an overlap for this region in the protein.

TV6-T λ 1 This peptide contains eight residues. The sequence of the first seven residues was determined by Edman degradation. C-terminal amino acid was deduced to be lysine because of the specificity of trypsin. The sequence was identical with that of the last five residues of TVh8 plus T λ 1 of protein Ha, except that Arg-108 of the latter was replaced by glycine in the present sample. This peptide overlaps the variable and the constant regions in the protein.

T λ 2 This peptide contains nineteen residues.

Fig. 3 Separation of chymotryptic peptides of the tryptic peptide TV1a on a column of Sephadex G-25 (1.5 $\times$ 90 cm). The chymotryptic digest dissolved in 0.5 ml of ammonia water was applied to the column equilibrated with 0.1 M NH_4HCO_3 . Conditions are the same as in Fig. 2.

The sequence of the first eighteen residues was determined by Edman degradation. The residue after eighteen cycles was applied to an amino acid analyzer after hydrolysis and lysine was detected. The repetitive yield of eighteen cycles of the degradation was on the average 92%. The sequence is identical with that of T λ 2 of protein Sh (25). Accordingly, the peptide was placed at the corresponding region of the present protein.

T λ 3a This peptide contains five residues. The sequence of the first four residues was determined by Edman degradation. The yield of the 3rd and 4th steps of the degradation was considerably low, but after modification with Braunitz's reagent (1), the yield of the 3rd and 4th steps was 80% and 77%, respectively. The C-terminal residue was deduced to be AE-cysteine because of the specificity of trypsin. The determined sequence was identical with that of T λ 3a of protein Sh (25).

T λ 3b This peptide containing a tryptophan residue consists of fifteen residues. The sequence of the first nine residues was determined by Edman degradation. The partial sequence and composition were identical with those of T λ 3b of protein Sh (25), and hence the two peptides were assumed to have the same sequence. Accordingly, the peptide was placed at the corresponding region of the present protein.

T λ 4 This peptide contains seven residues. The sequence of first four residues was determined by Edman degradation. The partial sequence and composition were identical with those of T λ 4 from Sh (25), and hence the two peptides were assumed to have the same sequence.

T λ 5 This peptide contains ten residues. The sequence of the first nine residues was determined by Edman degradation. The residue after nine cycles of the degradation was applied to an amino acid analyzer after hydrolysis and lysine was detected. The repetitive yield of nine cycles of the degradation was on average 81%. The sequence is identical with that of T λ 5 from protein Sh (25). Accordingly the peptide was placed at the corresponding region of the protein NIG-51.

T λ 6 This peptide contains five residues. N-terminal amino acid of the peptide was not detected by Edman degradation. It is tentatively assumed that pyrrolidone carboxylic acid occupied the N-terminal position. The amino acid composition was identical with that of peptide T λ 6 from protein Sh (25). Hence the two peptides were assumed to have the same se-

Table 3 Summary of Amino Acid Sequence of Tryptic Peptides of Protein NIG-51

Peptide	Sequence
TV1a	Glp-Ser-Val-Leu-Thr-Gln-Pro-Pro-Ser-Ala-Ser-Gly-Val-Pro-Gly-Gln-Ser-Val-Ile-Ile-Ser-AEC ← ← → → → → → → → → → → → → → → → → → →
TV1b	Ser-Gly-Ser-Ser-Ser-Asn-Ile-Gly-Arg → → → → → → → →
TV2	Asn-Thr-Val-Asn-Trp-Tyr-Gln-Gln-Val-Pro-Gly-Ala-Ala-Pro-Lys → → → → → → → → → → → → → → → →
TV3	Leu-Leu-Val-Tyr-Ser-Asn-Asn-Gln-Trp-Pro-Ser-Gly-Val-Pro-Asp-Arg → → → → → → → → → → → → → → → →
TV4	Phe-Ser-Gly-Ser-Lys → → → → →
TV5a	Ser-Gly-Thr-Ser-Ala-Ser-Leu-Ala-Ile-Ser-Gly-Leu-His-Ser-Glu-Asp-Glu-Ala-Asp-Tyr-Phe-AEC →
TV5b	Ala-Thr-Trp-Asp-Asp-Ser-Leu-Asp-Gly-Pro-Val-Phe-Gly-Gly-Gly-Thr-Lys → → → → → → → → → → → → → → → → →
TV5a-b	Ser-Gly-Thr-Ser-Ala-Ser-Leu-Ala-Ile-Ser-Gly-Leu-His-Ser-Glu-Asp-Glu-Ala-Asp-Tyr-Phe-AEC Ala-Thr-Trp-Asp-Asp (Asp, Thr, Ser, Pro, Gly ₄ , Val, Leu, Phe) Lys →
TV6-Tλ1	Leu-Thr-Val-Leu-Gly-Gln-Pro-Lys → → → → → → →
Tλ2	Ala-Ala-Pro-Ser-Val-Thr-Leu-Phe-Pro-Pro-Ser-Ser-Glu-Glu-Leu-Gln-Ala-Asn-Lys →
Tλ3a	Ala-Thr-Leu-Val-AEC → → → →
Tλ3b	Leu-Ile-Ser-Asp-Phe-Tyr-Pro-Gly-Ala (Val, Thr, Val, Ala, Trp) Lys → → → → → → → → → → → → → → → →
Tλ4	Ala-Asp-Ser-Ser (Pro, Val) Lys → → → → →
Tλ5	Ala-Gly-Val-Glu-Thr-Thr-Thr-Pro-Ser-Lys → → → → → → → → → → →
Tλ6	(Glx, Ser, Asx, Asx) Lys
Tλ7	Tyr-Ala-Ala-Ser-Ser-Tyr-Leu-Ser-Leu-Thr-Pro-Glu-Gln-Trp-Lys → → → → → → → → → → → → → → → →
Tλ8	Ser-His-Arg → → →
Tλ9a	Ser-Tyr-Ser-AEC → → → →
Tλ9b	(Glx, Val, Thr, His, Glx, Gly, Ser, Thr, Val, Glx) Lys
Tλ10	Thr-Val-Ala-Pro-Thr-Glu-AEC-Ser → → → → → → → →

(→) shows the sequence determined by the Edman degradation direct identification method; (→→) the sequence determined by the same procedure after modification with Braunitzer's reagent. (←) shows the sequence determined by carboxypeptidase digestions. AEC, aminoethylcysteine

quence, and Tλ6 was placed at the corresponding region of the protein NIG-51.

Tλ7 This peptide containing a tryptophan residue consists of fifteen residues. The sequence of the first thirteen residues was determined by Edman degradation, and the C-terminal amino acid was deduced to be lysine because of the specificity of trypsin. The determined

sequence was identical with that of Tλ7 from protein Sh (25).

Tλ8 This peptide was tripeptide. The N-terminal residue was determined by Edman degradation. The C-terminal residue was deduced to be arginine because of the specificity of trypsin. The sequence was identical with that of Tλ8 from protein Sh (25).

Tλ9a This peptide was tetrapeptide. The sequence of the first three residues was determined by Edman degradation. The C-terminal residue was deduced to be AE-cysteine because of the specificity of trypsin. The determined sequence was identical with that of Tλ9a from protein Sh (25). Accordingly, the peptide was placed at the corresponding region of the protein NIG-51.

Tλ9b This peptide contains eleven residues. The N-terminal residue was not detected by Edman degradation, because presumably pyrrolidone-carboxylic acid had been formed as an artifact by cyclization of a glutamine residue. The amino acid composition of the peptide was identical with that of Tλ9b from protein Sh (25), and hence the two peptides were assumed to have the same sequence. Accordingly the peptide was placed at the corresponding region of the protein NIG-51.

Tλ10 This peptide contains eight residues, and was sequenced throughout by Edman degradation. The repetitive yield of seven cycles of the degradation was on the average 80%. The residue after seven cycles was applied to an amino acid analyzer without hydrolysis and free serine was detected. The sequence was identical with that of Tλ10, the C-terminal peptide of protein Sh (25).

The above results of the sequence analysis of the tryptic peptides are summarized in Table 3.

The complete sequence of the variable region,

together with a partial sequence of the constant region, is given in Fig. 4.

DISCUSSION

The sequence presented here for the type L Bence Jones protein NIG-51 from a case of primary amyloidosis is characteristic of human λ light chains of the VλI subgroup. The protein contained 216 residues in total including two additional residues at positions 27b and 27c. The composition, complete or partial sequence data and alignment of tryptic peptides in the constant region (residues 108–213, Fig. 4) of the protein corresponded exactly to that of protein Sh (25). Thus, discussion will center on the variable region.

There are three tryptophan residues in the V region: they are Trp-33, Trp-52, and Trp-89, respectively. Of these, Trp-33 is common to all the λ chains so far tested, and Trp-89 is consistently present in the VλI subgroup. On the other hand, Trp-52 in the specimen is the first case at which amino acid replacement is demonstrated in the λ light chains. An arginine residue is generally present at the position in every instance so far reported. It may be assumed that the residue is in some way involved in the formation of the fibril complex through inter-chain hydrophobic interactions by itself, or in association with a number of different amino

Fig. 4 Amino acid sequence of the protein NIG-51. Tryptic peptides are indicated by the symbol T over their C-terminal residue. The symbol V is added for tryptic peptides in the variable region of the sequence (residues 0 through 108) and λ for those in the constant region (residues 109 through 213). The tryptic peptides in each half of the sequence are numbered consecutively.

Fig. 5 Comparison of the amino acid sequence of the variable region of proteins NIG-51 (this work), Ha (16), New (9), Vor (5), Vi (14), Sh (25), X (10), Del (6), and NIG-48 (22). To obtain maximum homology, some deletions (---) have been introduced into the figure.

acid residues present in a randomly coiled region (residues 41–62 in the V region).

Comparison with the sequence of other human λ chains (Fig. 5) indicates that the protein NIG-51 belongs to subgroup V λ I. The reasons are 1) positions 0–8, 10, 13–15, 17, 19–23, 27c, 29, 31, 33, 35, 38, 39, 41, 42, 44, 45, 47, 54, 55, 57–59, 61, 63–69, 71, 73, 75, 76, 81, 82, 84, 86, 87, 89, 90, 93, 97–102, and 105–107 of the variable region of the protein are occupied by the identical residues established in the sequence analyses of other proteins of subgroup V λ I; and 2) the protein has two additional residues at position 27, which are characteristic of subgroup V λ I.

It is of particular interest to note that there is an unusually high degree of homology in sequence between two proteins, NIG-51 and Protein Vor (5) (Fig. 5). The sequences of these proteins are identical from position 54 through 108 except for three residues at positions 77, 80, and 104. The region corresponds exactly to the half size of the V domain. In the stretch, the sequences are identical over 23 consecutive residues (regions 54–76, 81–104) including the third hypervariable region (around position 90). Thus, both the proteins are considered to have the same idiotypic specificity. In contrast to this finding, there are seventeen differences in the sequence in the other half of the domain (residues 0–53). The evidence strongly suggests that intrasubgroup recombination has occurred in the region around 54. Such a possibility has been suggested by Smith *et al.* (20), and was later demonstrated by one of us (T. S.) for the human κ chain (18).

We have already analyzed the structure of immunoglobulins (16, 17, 18, 19, 22) with the major objective of determining how the great variety of antibody sequences is generated. As for the number of variable genes, as described in a previous paper (17), if each subgroup is encoded by at least one germ line gene, much of the variability of immunoglobulins can be accounted for by assuming a relatively small number, for example $\sim 10^2$, of germ line genes. At present the number of subgroups is 12. Protein NIG-51 is also classified into subgroup V λ I, which has already been established. Therefore, it is likely that most published sequences of immunoglobulins fall into a small number of subgroups.

Smith and his coworkers described that subgroup V κ I was further subdivided into subsubgroups between which recombination occurs (20). On the basis of the subdivision, they proposed four V κ I genes against subgroup V κ I. If this subdivision is taken into consideration, one

hundred does not seem to be an unreasonable value for the number of variable genes for human immunoglobulins. Obviously more sequence data are needed to clarify the subdivisions.

The protein NIG-51 was isolated from the urine of a patient with primary amyloidosis. The sequence analysis showed that the protein NIG-51 was an intact λ light chain. Terry *et al.* demonstrated that the N-terminal sequence of urinary Bence Jones protein was identical with that of the amyloid fibril protein isolated from the small intestine of a patient with plasma cell dyscrasia and associated amyloidosis (23). They further showed that the predominant protein in the amyloid deposits of this patient was identical with the patient's urinary Bence Jones protein. The amino acid sequence of this κ Bence Jones protein (Tew) was determined by Putnam *et al.* (15). Unfortunately, we could not obtain the amyloid fibril protein from our patient. However, the protein NIG-51 is the first case in which the entire amino acid sequence of a λ Bence Jones protein from a case of primary amyloidosis was determined.

This work was supported in part by a grant from the Ministry of Education, Science and Culture, Japan, and by a research funds from Ito Science Foundation. We thank Dr H. Kasai and Dr T. Isobe for valuable discussions.

Received for publication 9 June 1980

REFERENCES

1. BRAUNITZER G., SCHRANK B. und RUHTUS A. (1970) Versuche zum vollständigen und automatischen Abbau von Peptiden nach der Quadrolmethode. *Hoppe-Seyler's Z. Physiol. Chem.* **351**, 1589–1590
2. COHEN A. S., CALKINS E. (1959) Electron microscopic observations on a fibrous component in amyloid of diverse origins. *Nature* **183**, 1202–1203
3. DAVIS B. J. (1964) Disc electrophoresis—II. Method and application to human serum proteins. *Ann. New York Acad. Sci.* **121**, 404–427
4. DUNKER A. K. and RUECKERT R. R. (1969) Observations on molecular weight determinations on polyacrylamide gel. *J. Biol. Chem.* **244**, 5074–5080
5. ENGELHARD M., HESS M. und HILSCHMANN N. (1974) Die Primärstruktur einer monoklonalen Immunglobulin-L-kette der Subgruppe I vom λ -Typ (Bence-Jones-protein Vor). *Hoppe-Seyler's Z. Physiol. Chem.* **355**, 85–88
6. EULITZ M. (1974) A new subgroup of human L-chains of the λ -type: Primary structure of Bence Jones Protein DEL. *Eur. J. Biochem.* **50**, 49–69

7. GLENNER G. G., TERRY W., HARADA M., ISERSKY C. and PAGE D. (1971) Amyloid fibril proteins: Proof of homology with immunoglobulin light chains by sequence analyses. *Science* **172**, 1150-1151
8. ISOBE T. and OKUYAMA T. (1978) The amino-acid sequence of S-100 protein (PAP I-b protein) and its relation to the calcium-binding proteins. *Eur. J. Biochem.* **89**, 379-388
9. LANGER V. B., STEINMETZ-KAYNE M. and HILSCHMANN N. (1968) Die vollständige Aminosäuresequenz des Bence-Jones-proteins New (λ -Typ): Subgruppen im variablen Teil bei Immunglobulin-L-ketten vom λ -Typ. *Hoppe-Seyler's Z. Physiol. Chem.* **349**, 945-951
10. MILSTEIN C., CLEGG J. B. and JARVIS J. M. (1968) Immunoglobulin λ -chains: The complete amino acid sequence of a Bence Jones protein. *Biochem. J.* **110**, 631-654
11. ORNSTEIN L. (1964) Disc electrophoresis—I. Background and theory. *Ann. New York Acad. Sci.* **121**, 321-349
12. OUCHTERLONY Ö. (1964) Diffusion-in-gel methods for immunological analysis II. In *Progress in Allergy* (ed. KALLÓS P. and WAKSMAN B. H.) S. Karger, Basel, Vol. 6, pp. 30-154
13. PENKE B., FERENCZI R. and KOVACS K. (1974) A new acid hydrolysis method for determining tryptophan in peptides and proteins. *Anal. Biochem.* **60**, 45-50
14. PONSTINGL V. H. and HILSCHMANN N. (1969) Vollständige Aminosäuresequenz einer λ -Kette der Subgruppe II (Bence-Jones-protein VII): Die Evolution als Ursache der Antikörperspezifität. *Hoppe-Seyler's Z. Physiol. Chem.* **350**, 1148-1152
15. PUTNAM F. W., WHITLEY E. J., PAUL C. and DAVIDSON J. N. (1973) Amino acid sequence of a κ Bence Jones protein from a case of primary amyloidosis. *Biochemistry* **12**, 3763-3780
16. SHINODA T., TITANI K. and PUTNAM F. W. (1970) Amino acid sequence of human λ -chains. I. Tryptic peptides of protein Ha. *J. Biol. Chem.* **245**, 4463-4474
17. SHINODA T. (1975) Comparative structural studies on the light chains of human immunoglobulins. I. Protein Ka with the Inv (3) allotypic marker. *J. Biochem.* **77**, 1277-1296
18. SHINODA T. (1973) Amino acid sequence of a human kappa type Bence-Jones protein. I. Tryptic peptides of protein Ni. *J. Biochem.* **73**, 417-431
19. SHINODA T. (1973) Amino acid sequence of a human kappa type Bence Jones protein. II. Chymotryptic peptides and sequence of protein Ni. *J. Biochem.* **73**, 433-446
20. SMITH G. P., HOOD L. and FITCH W. M. (1971) Antibody diversity. *Ann. Rev. Biochem.* **40**, 969-1012
21. SPACKMAN D. H., STEIN W. H. and MOORE S. (1958) Automatic recording apparatus for use in the chromatography of amino acids. *Anal. Chem.* **30**, 1190-1206
22. TAKAHASHI N., TAKAYASU T., ISOBE T., SHINODA T., OKUYAMA T. and SHIMIZU A. (1979) Comparative study on the structure of the light chains of human immunoglobulins. II. Assignment of a new subgroup. *J. Biochem.* **86**, 1523-1535
23. TERRY W. D., PAGE D. L., KIMURA S., ISOBE T., OSSERMAN E. F. and GLENNER G. G. (1973) Structural identity of Bence Jones and amyloid fibril proteins in a patient with plasma cell dyscrasia and amyloidosis. *J. Clin. Invest.* **52**, 1276-1281
24. TITANI K., WIKLER M. and PUTNAM F. W. (1970) The amino acid sequence of a λ type Bence-Jones protein. I. Tryptic peptides. *J. Biol. Chem.* **245**, 2142-2157
25. TITANI K., WIKLER M., SHINODA T. and PUTNAM F. W. (1970) The amino acid sequence of a λ -type Bence-Jones protein. III. The complete amino acid sequence and the location of the disulfide bridges. *J. Biol. Chem.* **245**, 2171-2176
26. VAN EERD J.-P. and TAKAHASHI K. (1976) Determination of the complete amino acid sequence of bovine cardiac troponin C. *Biochemistry* **15**, 1171-1180
27. WIKLER M. and PUTNAM F. W. (1970) Amino acid sequence of human λ chains. III. Tryptic peptides, chymotryptic peptides, and sequence of protein Bo. *J. Biol. Chem.* **245**, 4488-4507
28. WIKLER M., TITANI K. and PUTNAM F. W. (1970) The amino acid sequence of a λ type Bence Jones protein. II. Chymotryptic peptides. *J. Biol. Chem.* **245**, 2158-2170
29. YOKOYAMA S., OOBAYASHI A., TANABE O., SUGAWARA S., ARAKI E. and ICHISHIMA E. (1974) Production and some properties of a new type of acid carboxypeptidase of *Penicillium* molds. *Appl. Microbiol.* **27**, 953-960