Sheid, B., Bilik, E., and Biempica, L. (1970), *Arch. Biochem. Biophys. 140*, 437.

Shugart, L., Chastain, B. H., Novelli, G. D., and Stulberg, M. P. (1968), *Biochem. Biophys. Res. Commun. 31*, 404.

Tomkins, G. M., and Martin, D. W. (1970), *Annu. Rev. Genet. 4*, 91.

Vaughan, M. H., Soeiro, R., Warner, J. R., and Darnell, J. E. (1967), *Proc. Nat. Acad. Sci. U. S. 58*, 1527.

Amino Acid Sequence of a κ Bence Jones Protein from a Case of Primary Amyloidosis†

Frank W. Putnam,* E. J. Whitley, Jr.,‡ C. Paul,§ and J. N. Davidson¶

ABSTRACT: Amino acid sequence analysis has been done on a κ Bence Jones protein (Tew) from a case of primary amyloidosis with the objective of determining the sequence of the variable region. Twenty-two tryptic peptides accounting for 182 residues were isolated and were completely or partially sequenced. Chymotryptic digestion yielded 32 peptides which supplied many overlaps. Sequenator analysis was performed for the first 42 residues of the amino terminus. From the combined data the sequence of the variable region (residues 1–108) was deduced. The composition, partial sequence data, and alignment of the peptides of the constant region (residues 109–214) correspond exactly to the sequence established for human κ light chains of the same allotype. The V region includes five extra residues (30a–30e) and is characteristic of the κII subgroup. A computer analysis of the sequence of the Tew Bence Jones protein in comparison with other human κ light chains was undertaken to establish quantitative criteria for subgroup classification. In terms of minimum nucleotide mutations the Tew protein differs from other subgroup κII proteins by an average of only 0.2 base/amino acid residue position, whereas other human κ chains differ from Tew by an average of 0.4–0.6 base/position. The Bence Jones protein and the tissue amyloid protein from this patient appear to be identical in primary structure as indicated by identity in the amino-terminal sequence for 27 residues and similarity in peptide maps, amino acid composition, and other properties.

Since our report of the first complete amino acid sequence of a human κ Bence Jones protein (Putnam *et al.*, 1966), almost 5000 residue positions have been reported for other human light chains; yet, no two human light chains have yet proved to be identical in sequence. Each of the two types of light chains, κ and λ, has been shown to be divided almost precisely in half into a variable (V) region comprising the NH_2-terminal half and a constant (C) region comprising the COOH-terminal half (Putnam, 1969). The C region containing residues 109–214 is identical in sequence for all human κ light chains except for an inherited variation at position 191, where either valine or leucine may be present according to the allotype. In contrast, the V regions of any two human κ chains differ by a minimum of about ten and a maximum of about 60 residues. Human λ light chains conform to a similar principle of structure, as do the light chains of most other species. Nonetheless, some human κ chains are more related in amino acid sequence in their V regions than are others. Such structurally related light chains are said to belong to the same subgroup. Three such subgroups have been proposed for human kappa light chains, *i.e.*, κI, κII, and κIII (Putnam, 1969; Milstein, 1969; Hilschmann, 1969).[1] However, the subgroups have been defined arbitrarily and are based largely on short NH_2-terminal sequences rather than on complete sequences of the V region. Since the existence of subgroups is a fundamental postulate of several theories of antibody diversity, a more exacting analysis is needed of the subgroup classification of a series of light chains that have been completely or nearly completely sequenced. This article presents the proof for the amino acid sequence of a human Bence Jones protein (Tew) which belongs to the κII subgroup, and discusses the sequence relationship of this light chain to other κ light chains. This Bence Jones protein was isolated from the urine of a patient who died of primary amyloidosis (Case 14 of Osserman *et al.*, 1964), and the sequence of the Bence Jones protein and of the amyloid protein isolated from his tissues is identical as far as the latter could be determined (Terry *et al.*, 1973).

Experimental Procedure

Bence Jones protein Tew was precipitated from the urine of the patient by adjustment to 60% ammonium sulfate and was purified by use of a DEAE-Sephadex A-50 column (Bernier

† From the Department of Zoology, Indiana University, Bloomington, Indiana 47401. *Received April 24, 1973.* This work was supported in part by grants from the National Institutes of Health (CA 08497), the American Cancer Society (NP-10), and the Damon Runyon Fund for Cancer Research (DRG 1134). The manual sequence data are documented in the M.S. thesis of E. J. W. (1969).

‡ Present address: Department of Biochemistry, University of Minnesota, Saint Paul, Minn. 55101.

§ Present address: Laboratoire de Chimie générale, Université Libre de Bruxelles, Belgium.

¶ Present address: Biology Department, Harvard University, Cambridge, Mass. 02138. National Institutes of Health trainee supported by GM-00036.

[1] For nomenclature of human κ chain subgroups see the Discussion and also statements in *Bull. W. H. O.* [(1969), *41*, 975] and in *Biochemistry* [(1972), *11*, 3311].

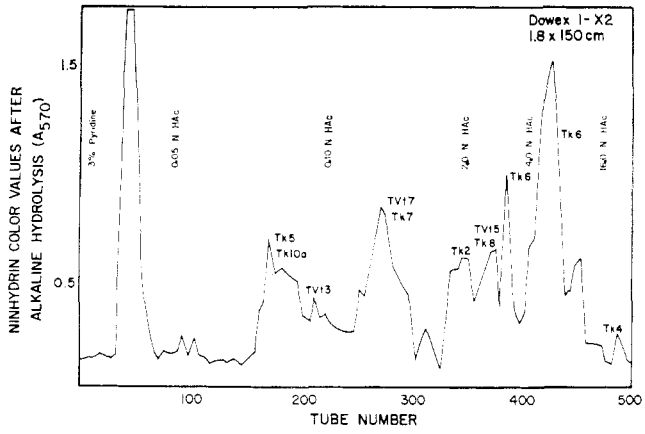


FIGURE 1: Chromatographic pattern of the peptides from a tryptic digest of the Tew Bence Jones protein on a 1.8 × 150 cm column of Dowex 1-X2. The column was developed at 37° at a flow rate of 70 ml/hr with a pyridine–acetic acid gradient as described in the text. Fractions of 10 ml were collected and monitored by ninhydrin analysis at 570 nm after alkaline hydrolysis.

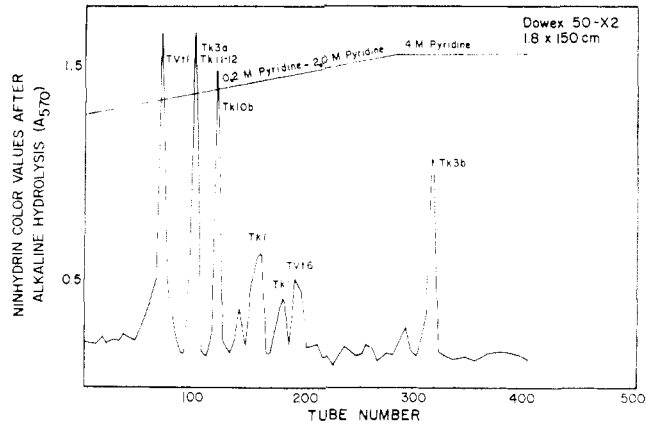


FIGURE 2: Chromatographic pattern on a Dowex 50-X2 column of the basic tryptic peptides in the first peak of Figure 1. The column of 1.8 × 150 cm was developed at 50° at a flow rate of 75 ml/hr with a gradient from 0.2 M pyridine to 2.0 M pyridine. At the completion of the gradient 4 M pyridine was passed through the column to elute any remaining peptides. Fractions of 10 ml were collected and monitored as in Figure 1.

and Putnam, 1964). The homogeneity was established by ultracentrifugation, starch gel electrophoresis, immunodiffusion, and immunoelectrophoresis (Bernier and Putnam, 1964). The latter two procedures showed the protein to be of the κ type. The allotype of Protein Tew was determined as Inv (3) by Dr. Arthur G. Steinberg of Case Western Reserve University. Inv (3) is equivalent to the former notation Inv (a − b +). The general methods for the preparation, isolation, characterization, and manual sequence determination of the peptides and the sources of enzymes, reagents, and ion-exchange resins are the same as those described previously (Titani *et al.*, 1969, 1970; Shinoda *et al.*, 1970). Further details are given below. Prior to enzymatic digestion the protein was reduced with mercaptoethanol in 7 M guanidine hydrochloride and then alkylated either with ethylenimine or iodoacetamide (Titani *et al.*, 1970). The aminoethylated protein was used for the tryptic digest to increase the points of cleavage; carboxymethylated protein was used for one chymotryptic digest and aminoethylated protein for another.

Amino Acid Analysis and Sequence Determination. For amino acid analysis samples were hydrolyzed for 22 hr with 6 N HCl at 110° in an evacuated sealed tube and analyzed with the Beckman Model 120 automatic amino acid analyzer equipped with high sensitivity cuvettes and recorders. Manual sequence determination of peptides was done as previously described by use of the Edman degradation (Titani *et al.*, 1969, 1970; Shinoda *et al.*, 1970). Digestion with leucine aminopeptidase or with carboxypeptidase A was also used for end-group analysis and to determine amide positions. Small peptides were sequenced completely by the Edman degradation. Larger tryptic peptides were digested further with chymotrypsin, pepsin, Pronase, or papain to yield smaller subpeptides; the sequence of the original peptide was determined both by Edman degradation and by the overlapping of the subpeptides. In a few cases the amino end groups of peptides were determined by the dansyl Edman method (Gray, 1967) and the dansyl derivatives were characterized by the procedure of Woods and Wang (1967). In later work after the Beckman Model 890 Sequencer became available, the amino terminus of the whole protein was sequenced automatically, twice for 27 cycles and once for 42 cycles. The phenylthiohydantoin (Pth) derivatives were identified by gas chromatography with two columns (Pisano and Bronzent, 1969) with verification of the regenerated amino acid as needed. For automatic sequence analysis procedures similar to those of Hermodson *et al.* (1972a) were followed.

Tryptic Digest. The digest of 500 mg (21 μmol) of the aminoethylated protein with trypsin (treated with L-tosylamido-2-phenylethyl chloromethyl ketone) was prepared at 37°, pH 8.0, in the pH-Stat. Trypsin was added at a weight ratio of enzyme to substrate of 2% in a final volume of 250 ml, and the reaction was stopped after 12 hr by adjustment to pH 4.0 with 6 N acetic acid.

Chymotryptic Digest. Two chymotryptic digests were prepared, the first with 200 mg of the aminoethylated protein and the second with 450 mg of the carboxymethylated protein. The enzyme was added to give a final concentration of 2% by weight of substrate; digestion was done at pH 8.0 with the pH-Stat for 3 hr. The solution was lyophilized and dissolved in the appropriate buffer for column chromatography.

Results

Isolation of the Tryptic Peptides. The tryptic digest was first fractionated on a Dowex 1-X2 column which yielded the separation shown in Figure 1. The column was developed with a starting buffer of 3% pyridine and a gradient from 0.05 to 16 N acetic acid similar to that used in our previous work (Shinoda *et al.*, 1970). The fractions were pooled and examined for purity by high-voltage electrophoresis. Of the 13 peptides shown in Figure 1, only T_K6 was pure enough to initiate sequence determination. The other peaks represented a mixture of peptides and had to be purified further on a column of Dowex 50-X2, and in some cases by additional purification by paper electrophoresis.

The first peak shown in Figure 1 represented the basic peptides which were not retained at all on Dowex 1. The pool from this peak was lyophilized and fractionated on a column of Dowex 50-X2 (Figure 2). With the exception of the second peak, the peptides were recovered in homogeneous form, suitable for sequence analysis. The second peak was subjected to paper electrophoresis to separate the mixture into two peptides.

The procedure described above is similar to that used in our previous separations of the tryptic peptides of light chains.

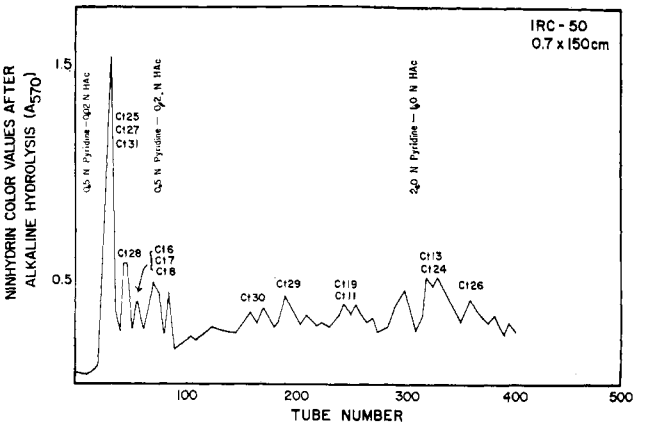


FIGURE 3: Chromatographic pattern on IRC-50 of the peptides from the first chymotryptic digest of Bence Jones protein Tew eluted with a gradient of pyridine–acetate buffers. Both the mixing chamber (800 ml) and the reservoir (800 ml) contained 0.05 N pyridine–0.02 N acetic acid, but the buffer in the reservoir was replaced with 0.5 N pyridine–0.2 N acetic acid at tube 70 and with 2 N pyridine–1 N acetic acid at tube 300, respectively (Titani *et al.*, 1969). The fraction size was 10 ml.

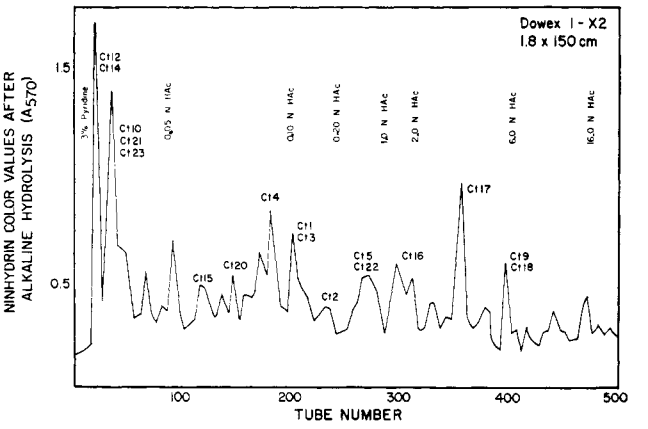


FIGURE 4: Chromatographic pattern on Dowex 1-X2 of the peptides from the second chymotryptic digest of Bence Jones protein Tew eluted with a gradient from 3% pyridine to 16 N acetic acid.

However, because the tryptic peptides of the V region may differ greatly from one κ chain to another, the elution pattern will vary and likewise the separation procedures needed for final purification of the peptides.

Nomenclature of Tryptic Peptides. In Figures 1 and 2 the peptides are designated according to their position in the sequence rather than in the order of elution. In accord with our previous nomenclature (Titani *et al.*, 1969), the capital letter T refers to the enzyme trypsin, V to the variable region, κ to the constant region, and the lower case t to the protein Tew. The symbol κ is used to indicate that identical tryptic peptides are expected for the constant region of all human κ chains except for hereditary allotypic differences at position 191 (Putnam, 1969). Certain κ proteins such as Tew (and also certain λ proteins) have extra residues at or about position 30. These are designated 30a, 30b, etc. in order to keep the numbering in the constant region invariant. On the other hand, the chymotryptic peptides are numbered consecutively throughout the length of the sequence since the number, variety, and yield of chymotryptic peptides depend much more on the conditions of digestion than for the tryptic peptides. Also, since the half-cystine residues were aminoethylated in the tryptic digest, additional peptides ending in aminoethylcysteine were obtained; these are given the same number as for the tryptic peptide expected from the unsubstituted protein plus the symbol "a," whereas the remaining portion of the peptide has the same number, plus the symbol "b."

Isolation of the Chymotryptic Peptides. The first chymotryptic digest of carboxymethylated protein was separated on IRC-50 with a smaller column (0.7 × 150 cm) because of the smaller sample size (200 mg). None of the peaks shown in Figure 3 contained peptides sufficiently pure for amino acid sequence determination, so additional purification by paper electrophoresis was necessary.

The second chymotryptic digest (450 mg) was placed on a column of Dowex 1-X2 with the same gradient used for the tryptic digest in Figure 1. Only two of the peptides identified in Figure 4 were isolated in homogeneous form. The other peaks were resubmitted to column chromatography on a Dowex 50-X2 column for further purification. The method of purification was similar to that for the tryptic peptides in Figure 2, but a Technicon Autoanalyzer was used for detection of the peptides rather than the manual procedure.

Amino Acid Composition of the Tew Bence Jones Protein and of Its Tryptic and Chymotryptic Peptides. The amino acid composition of the tryptic peptides is given in Table I and that of the chymotryptic peptides in Table II. For convenience and later reference the peptides are listed in sequential order from the amino terminus to the carboxyl terminus and the position in the sequence is given. The amino acid composition of the whole protein is given in Table III and also the composition of the tissue amyloid protein from the same patient as recalculated from data of Terry *et al.* (1973). On the basis of the content of lysine (11 residues), arginine (8 residues), and Ae-cysteine (5 residues), 25 tryptic peptides would be expected if all susceptible bonds were fully cleaved. However, only 17 nonoverlapping tryptic peptides were obtained plus 3 overlapping peptides. The close similarity in amino acid composition of the Bence Jones protein and the tissue amyloid protein is evident.

A special problem in this work was the small amount of protein available for sequence analysis (about 1.5 g) and the impossibility of obtaining more protein because of the death of the patient. However, because of the unusual features of this case the work was undertaken. The shortage of protein prevented the rigorous purification of some of the peptides lest they be lost with corresponding gaps in the sequence. As a result, the analyses of some of the tryptic and chymotryptic peptides are off in some cases by several tenths of a residue. In certain cases pointed out in the text, a combination of analytical and sequence data from the chymotryptic peptides related to a tryptic peptide is required to remove the ambiguity about the latter, *e.g.*, Ct20-Ct23 are required to establish the presence of only one proline residue in TVt6 although the analysis of the tryptic peptide suggests two.

Amino Acid Sequence Analysis

The emphasis in this work was placed on the V region of the Tew κ chain since the C region sequence of the κ Bence Jones protein Ag (Putnam *et al.*, 1966) has since been verified by independent study of at least 10 human κ chains in several laboratories (Hilschmann, 1967; Dreyer *et al.*, 1967; Milstein, 1969; Gottlieb *et al.*, 1970; Watanabe and Hilschmann, 1970; Köhler *et al.*, 1970). To determine the V region sequence an attempt was made to isolate all the tryptic peptides of the Tew κ chain, to sequence all tryptic peptides not identified with the C region, and to cover undetermined regions through

TABLE I: Amino Acid Composition[a] of the Tryptic Peptides of Bence Jones Protein Tew.

	TVt1	TVt2	TVt3	TVt3–TVt4	TVt4	TVt5[f]	TVt6	TVt7
Lysine		(1)		0.94 (1)	0.94 (1)			1.02 (1)
Histidine		(1)						
Arginine	1.22 (1)	(1)	0.84 (1)	1.00 (1)		0.69 (1)	1.05 (1)	
Aminoethylcysteine	0.91 (1)					0.63 (1)		
Aspartic acid	0.84 (1)	(4)	0.93 (1)	1.81 (2)	0.96 (1)	1.01 (1)		
Threonine	1.92 (2)			1.93 (2)	1.70 (2)		1.84 (2)	
Serine	3.50 (4)	(6)	0.89 (1)	3.62 (4)	2.46 (3)	0.88 (1)		
Glutamic acid	2.04 (2)	(4)				2.02 (2)	2.82 (3)	0.93 (1)
Proline	4.11 (4)	(2)	1.30 (1)	1.30 (1)			1.62 (1)	
Glycine	1.02 (1)	(2)	1.10 (1)	4.30 (4)	3.00 (3)	1.10 (1)	1.85 (2)	
Alanine	1.01 (1)	(1)	1.00 (1)	0.88 (1)		1.00 (1)	1.94 (2)	
Valine	1.62 (2)[c]		1.10 (1)	0.96 (1)		2.75 (3)		
Methionine	0.77 (1)						0.80 (1)	
Isoleucine	1.57 (2)[c]	(1)				0.86 (1)	0.98 (1)	1.00 (1)
Leucine	2.00 (2)	(7)		1.00 (1)	1.20 (1)		1.00 (1)	1.03 (1)
Tyrosine		(3)				1.90 (2)		
Phenylalanine		(1)		1.93 (2)	2.02 (2)		1.04 (1)	
Tryptophan		(1)						
Total residues	24	35[e]	7	20	13	14	15	4
Per cent yield[b]	25		22	3	14	4	8	35
Residue no.	1–24	25–54	55–61	55–74	62–74	75–88	89–103	104–107

	TVt7–Tκ1	Tκ1	Tκ2	Tκ3a	Tκ3b	Tκ4	Tκ5
Lysine	0.81 (1)		1.13 (1)			0.95 (1)	0.81 (1)
Histidine							
Arginine	0.87 (1)	1.00 (1)			1.27 (1)		
Aminoethylcysteine				0.75 (1)		0.22	
Aspartic acid			0.96 (1)		1.98 (2)		
Threonine			0.73 (1)	0.80 (1)			
Serine			1.65 (2)	2.04 (2)			
Glutamic acid	0.91 (1)		1.95 (2)			1.00 (1)	1.05 (1)
Proline			3.30 (3)		1.06 (1)		
Glycine				0.93 (1)		0.24	
Alanine			1.60 (2)	1.00 (1)		0.72 (1)	
Valine			1.60 (2)	1.68 (2)			1.00 (1)
Methionine							
Isoleucine	1.00 (1)		0.85 (1)				
Leucine	1.20 (1)		1.00 (1)		2.00 (2)		
Tyrosine					0.87 (1)		
Phenylalanine			1.70 (2)		1.17 (1)		
Tryptophan							+ (1)
Total residues	5	1	18	8	8	3	4
Per cent yield[b]	7	38	3	32	10	17	42
Residue no.	104–108	108	109–126	127–134	135–142	143–145	146–149

	Tκ6	Tκ7	Tκ8	Tκ9	Tκ10a	Tκ10b	Tκ11-12
Lysine	0.66 (1)	0.95 (1)	1.27 (1)	(1)		0.88 (1)	
Histidine				(1)		0.86 (1)	
Arginine							0.74 (1)
Aminoethylcysteine					0.68 (1)		(1)[d]
Aspartic acid	3.76 (4)	0.98 (1)	0.97 (1)				1.05 (1)
Threonine	0.91 (1)	2.90 (3)				2.00 (2)	
Serine	3.30 (4)	4.82 (5)				2.31 (2)	0.85 (1)
Glutamic acid	5.00 (5)		1.00 (1)			2.00 (2)	1.01 (1)
Proline						1.50 (1)	
Glycine	1.02 (1)					0.95 (1)	1.26 (1)
Alanine	1.02 (1)		0.89 (1)		1.00 (1)		
Valine	1.82 (2)				1.00 (1)	2.00 (2)	
Methionine							

TABLE I (*Continued*)

	Tκ6	Tκ7	Tκ8	Tκ9	Tκ10a	Tκ10b	Tκ11-12
Isoleucine							
Leucine	1.13 (1)	3.00 (3)				1.00 (1)	
Tyrosine		0.80 (1)	0.98 (1)		1.00 (1)		
Phenylalanine							1.00 (1)
Tryptophan							
Total residues	20	14	5	2	4	13	7
Per cent yield[b]	30	8	24		10	21	10
Residue no.	150–169	170–183	184–188	189–190	191–194	195–207	208–214

[a] Values are given in residues per molecule. Impurities are listed only if greater than 10% of the unit residue value. The assumed integral levels are given in parentheses. Peptide TVt2 was not isolated; the values for it given in parentheses are based on the sequence from the chymotryptic peptides and by sequenator analysis of the whole protein. Peptide Tκ9 is the basic dipeptide His-Lys; this was identified by its characteristic position and staining reactions in fingerprints of the tryptic digest (Easley and Putnam, 1966) but was lost during chromatography. [b] Uncorrected yield calculated on the amount of material taken for the tryptic digest. [c] Low values for valine and isoleucine were obtained in peptide TVt1 because the Ile–Val bond at positions 2–3 is resistant to acid hydrolysis for 24 hr. [d] *S*-Aminoethylcysteine was clearly present but too low to evaluate accurately. See Ct32 in Table II. [e] Although the residue numbers of TVt2 are given as 25–54 (only 30 residues), the total number of residues is 35 because residues 30a–30e are included in this segment. [f] The composition of the tripeptide TVt5′ is: Ile, 1.0; Ser, 1.0; Arg, 1.0.

analysis of chymotryptic peptides and with the sequencer. Chymotryptic peptides not identified with undetermined regions were studied only in sufficient detail to establish overlaps of the tryptic peptides and as needed for verification purposes. Twenty-two tryptic peptides were isolated; the amino acid analyses are given in Table I, and the sequence data are given in a later summary table. Thirty-three chymotryptic peptides were isolated, the amino acid analyses are given in Table II, and sequence data for a few peptides are given in a later summary table. Detailed analytical data on Edman degradations and on the composition of subpeptides are given in the dissertation by Whitley (1969). The discussion of results for each peptide refers to Figure 5 which assembles all the data, indicates all peptides and subpeptides studied, and identifies chemical and enzymatic degradations by characteristic symbols.

Sequence Analysis of Tryptic Peptides

Peptide TVt1. Three steps of the Edman degradation gave the amino-terminal sequence Asx-Ile-Val (Table IV and Figure 5). Digestion with pepsin yielded four subpeptides (P1–P4), and three (C1–C3) were obtained with chymotrypsin. Peptides TVt1, TVt1-P1, and TVt1-C1 showed low recovery values for isoleucine and valine after hydrolysis for 22 hr because of the Ile–Val bond at positions 2–3. The latter two peptides had identical composition and contained methionine as the fourth residue; an identical tetrapeptide (CNBr I) was obtained by cleavage of the whole protein with CNBr in 70% formic acid. TVt1-P1 was acidic on paper electrophoresis at pH 6.5, indicating that the Asx must be aspartic acid.

Four steps of Edman degradation showed the amino-terminal sequence of TVt1-P2 to be Thr-Glx-Ser-Pro. Digestion of TVt1-P2 with carboxypeptidase A for 15 min released leucine, 1.0 mol/mol and serine, 0.2 mol/mol. This indicated the sequence of Ser-Leu as carboxy terminal but did not exclude the possibility of Ser-Leu-Leu. However, the sequence was made clear by the two chymotryptic subpeptides TVt1-C2 (Thr,Glx,Ser,Pro,Leu) and TVt-C3, which Edman degradation showed to be Ser-Leu. Digestion of TVt1-P2 with leucine aminopeptidase did not give free glutamic acid but neither was glutamine seen because it is eluted in the same position as serine. Paper electrophoresis showed TVtl-P2 to be neutral; this indicated the presence of glutamine, which was confirmed later by sequencer analysis.

The amino terminus of TVt1-P3 was established for six steps by Edman degradation. Digestion of TVt1-P3 with carboxypeptidase A for 6 hr released only serine. The carboxyl-terminal sequence was assumed to be Pro-Ala-Ser because serine would not have been split off if proline was the penultimate residue. This was later verified by sequencer analysis of the whole protein which also confirmed that the Glx residue was glutamic acid, as was indicated by the fact that TVt1-P3 was acidic on electrophoresis.

Three rounds of Edman degradation gave the sequence of TVt1-P4 as Ile-Ser-Cys-Arg. The half-cystine was present here as the basic residue aminoethylcysteine, but the Edman degradation showed that arginine must be at the C terminus. Thus, trypsin did not cleave the bond between these two basic residues. As described later, the order of the subpeptides TVt1-P2 and TVt1-P3 was established from the chymotryptic peptides of the whole protein and by sequenator analysis of the whole protein.

Peptide TVt2. This hypothetical tryptic peptide containing 35 residues (positions 25–54) was not isolated; it was probably lost because of its large size. The sequence in this region was extended by sequenator analysis of the whole protein and was determined in part from the chymotryptic peptides of the whole protein (see later).

Peptide TVt3. Five steps of the Edman degradation indicated the N-terminal sequence to be Ala-Ser-Gly-Val-Pro. Because of the specificity of trypsin the rest of the sequence must be Asx-Arg. TVt3 behaved as a neutral peptide in electrophoresis at pH 6.5; this showed that the Asx must be Asp, giving the complete sequence as Ala-Ser-Gly-Val-Pro-Asp-Arg (Figure 5).

Peptide TVt4. Two steps of the Edman degradation showed the N-terminal sequence to be Phe-Ser. Additional steps gave unclear results because of the large number of serine and glycine residues in succession. Digestion of TVt4 with pepsin yielded three subpeptides (P1–P3), and digestion with papain gave four (Pap1–Pap4). The partial sequence of TVt4 was assembled from these (Table V) and was completed through the chymotryptic peptides of the whole protein (Figure 5).

TVt4-P1 was the amino-terminal subpeptide; attempts at

TABLE II: Amino Acid Composition[a] of Chymotryptic Peptides of Bence Jones Protein Tew.

	Ct1	Ct2	Ct3	Ct4	Ct4′	Ct5	Ct6
Lysine							
Histidine						0.72 (1)	
Arginine				0.93 (1)			
Aminoethylcysteine				0.77 (1)[b]			
Aspartic acid	0.90 (1)					1.08 (1)	2.00 (2)
Threonine	0.90 (1)	1.00 (1)	0.89 (1)	0.95 (1)			
Serine	1.00 (1)	1.74 (2)	2.52 (3)	4.81 (5)	0.96 (1)	0.96 (1)	0.42
Glutamic acid	1.05 (1)	1.10 (1)	1.03 (1)	1.98 (2)			0.30
Proline	1.18 (1)	3.28 (3)	2.83 (3)	2.80 (3)			
Glycine		1.00 (1)	1.00 (1)	1.03 (1)		1.06 (1)	0.42
Alanine		1.00 (1)	1.00 (1)	1.00 (1)			
Valine	0.63 (1)	0.93 (1)	1.01 (1)	0.96 (1)			
Methionine	0.30 (1)						
Isoleucine	0.56 (1)	1.00 (1)	0.93 (1)	1.00 (1)			
Leucine	1.00 (1)	1.00 (1)	1.01 (1)	1.07 (1)	2.00 (2)		1.00 (1)
Tyrosine							0.97 (1)
Phenylalanine						1.00 (1)	
Tryptophan							+ (1)
Total residues	9	12	13	18	3	5	5
Residue no.	1–9	10–21	10–22	10–27	28–30	30a–30e	31–35
N terminus	Asp					His	

	Ct16	Ct17	Ct18	Ct19	Ct20	Ct21	Ct22	Ct23
Lysine								
Histidine								
Arginine	1.00 (1)		0.99 (1)	0.72 (1)			1.06 (1)	0.96 (1)
Aminoethylcysteine					0.93 (1)			
Aspartic acid	1.09 (1)	1.13 (1)	1.02 (1)	1.12 (1)				
Threonine		0.90 (1)				0.72 (1)	0.89 (1)	0.92 (1)
Serine	0.91 (1)	2.68 (3)	0.87 (1)	0.80 (1)				
Glutamic acid			2.00 (2)	1.96 (2)	1.03 (1)	0.12 (1)	1.02 (1)	0.92 (1)
Proline	1.03 (1)					0.85 (1)		
Glycine	1.18 (1)	3.00 (3)	0.98 (1)	1.04 (1)			1.98 (2)	1.00 (1)
Alanine	1.00 (1)		1.00 (1)	1.00 (1)	0.90 (1)	1.00 (1)		
Valine	1.03 (1)		2.90 (3)	2.88 (3)				
Methionine					0.73 (1)			
Isoleucine			1.01 (1)			0.78 (1)		
Leucine					1.00 (1)	0.34	1.00 (1)	1.04 (1)
Tyrosine			0.78 (1)	0.59 (1)				
Phenylalanine	0.94 (1)	0.89 (1)				0.73 (1)		
Tryptophan								
Total residues	8	9	12	11	5	6	6	5
Residue no.	55–62	63–71	75–86	76–86	88–92	93–98	99–104	100–104
N terminus		Ser				Phe		

[a] Values are given in residues per molecule. Impurities are listed only if greater than 10%. The assumed values are given in parentheses. [b] Peptide Ct4 was recovered from the chymotryptic digest in which the alkylating agent was iodoacetic acid. The

Edman degradation were unsuccessful. Because TVt4-P1 was acidic on electrophoresis at pH 6.5, the Asx must be aspartic acid. TVt4-P2 represented the remainder of TVt4. In accord with the specificity of pepsin phenylalanine must be carboxyl terminal in P1 and amino terminal in P-2, for P-3 lacked this residue.

The four papain subpeptides together account for all of TVt4. Except for Pap 3, Edman data were not obtained because of the small amount of the subpeptides. However, from inspection of the composition of the pepsin and papain subpeptides, as aligned in Table V, the COOH-terminal sequence of TVt4 could be deduced as Thr-Asp-Phe-Thr-Leu-Lys. The acidic mobility of TVt4-Pap 3 in electrophoresis at pH 6.5 confirmed the presence of aspartic acid rather than asparagine. The repetitious Ser-Gly sequence in Tvt4 was determined from the chymotryptic peptide Ct17 (see later).

Peptide TVt3–TVt4. Although obtained in only a 3% yield, this peptide provides evidence for the overlapping of TVt3 and TVt4. This tryptic peptide is attributed to incomplete cleavage of the Arg–Phe bond at positions 61–62 because of aspartic acid at residue 60.

Peptide TVt5. This peptide, also obtained in low yield (4%),

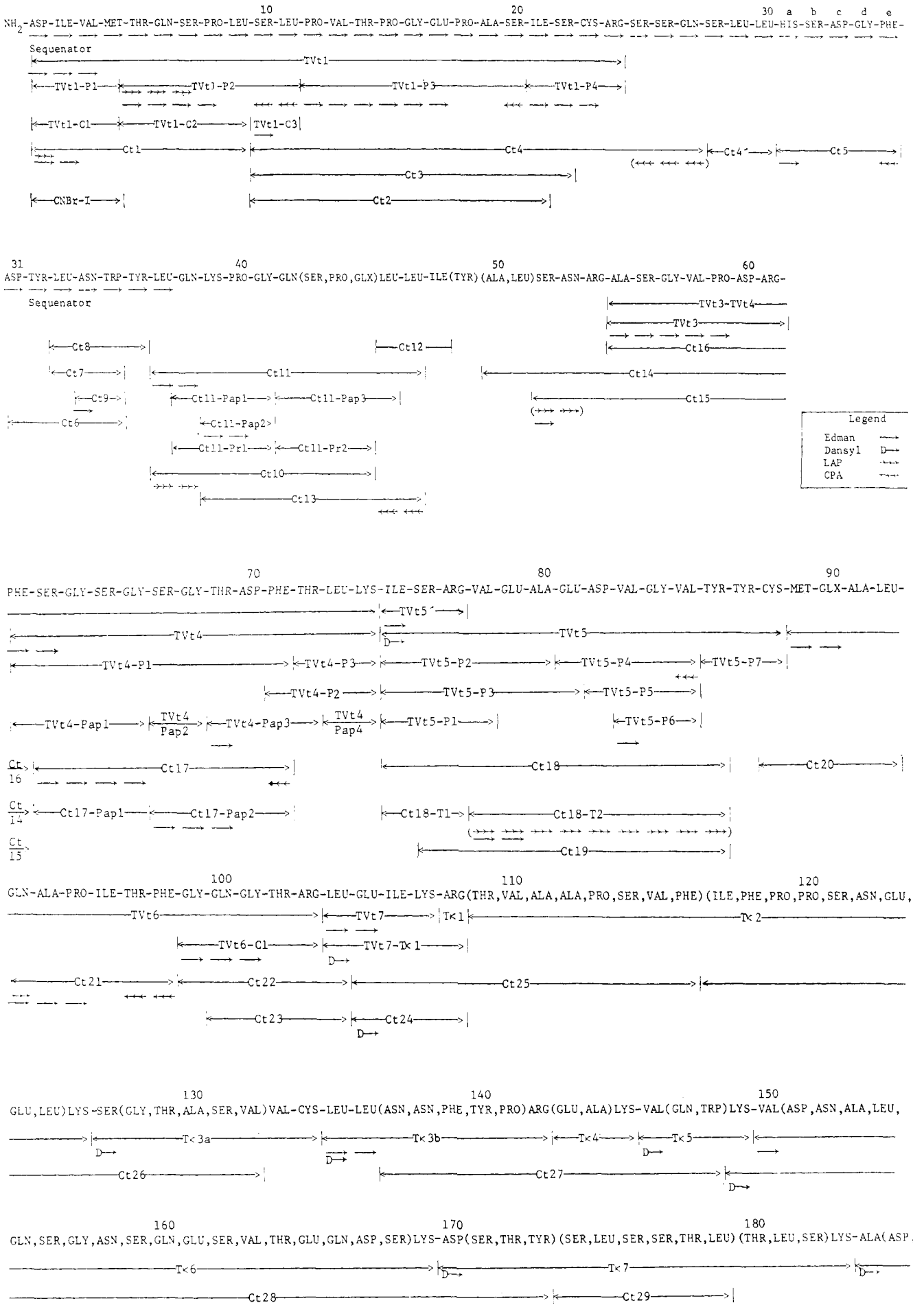
Ct7	Ct8	Ct9	Ct10	Ct11	Ct12	Ct13	Ct14	Ct15
			1.06 (1)	0.85 (1)		0.84 (1)		
							1.92 (2)	1.91 (2)
0.92 (1)	1.00 (1)	1.00 (1)					2.06 (2)	2.09 (2)
			1.00 (1)	1.14 (1)		1.26 (1)	1.88 (2)	1.91 (2)
			3.00 (3)	2.74 (3)		2.25 (2)		
			2.31 (2)	1.80 (2)		1.90 (2)	1.00 (1)	0.91 (1)
			1.00 (1)	1.00 (1)		1.00 (1)	1.06 (1)	1.09 (1)
				0.21			2.00 (2)	1.00 (1)
							1.03 (1)	1.09 (1)
					0.83 (1)			
1.00 (1)	0.64 (1)		1.00 (1)	2.71 (3)	2.00 (2)	1.95 (2)	0.86 (1)	
	0.47 (1)							
							0.97 (1)	0.91 (1)
+ (1)	+ (1)	+ (1)						
3	4	2	9	11	3	9	13	11
33–35	33–36	34–35	37–45	37–47	46–48	39–47	50–62	52–62
		Asn	Leu	Leu				Ser

Ct24	Ct25	Ct26	Ct27	Ct28	Ct29	Ct30	Ct31	Ct32
0.76 (1)	1.07 (1)	1.04 (1)	0.91 (1)	1.89 (2)			1.03 (1)	
						0.90 (1)		
1.19 (1)	1.11 (1)		1.10 (1)					0.87 (1)
						1.02 (1)		0.96 (1)
		1.17 (1)	1.68 (2)	4.80 (5)				0.92 (1)
	0.58 (1)	1.01 (1)		1.89 (2)	0.94 (1)	0.92 (1)	0.98 (1)	
	0.79 (1)	2.84 (3)	0.40	4.57 (5)	2.80 (3)		3.11 (3)	
1.00 (1)	0.93 (1)	2.14 (2)	1.68 (2)	4.88 (5)	0.16	1.99 (2)		0.99 (1)
	1.04 (1)	2.19 (2)	0.67 (1)				0.95 (1)	
	0.14	1.01 (1)	0.40	1.05 (1)		1.02 (1)		1.00 (1)
	2.28 (2)	1.00 (1)	1.00 (1)	1.05 (1)	0.14	1.00 (1)		
	2.24 (2)	1.34 (1)	0.91 (1)	2.02 (2)		0.98 (1)	1.00 (1)	
0.81 (1)	1.00 (1)	0.73 (1)						
	0.11	0.99 (1)	0.40	1.00 (1)	2.00 (2)	0.98 (1)		
			0.81 (1)	0.78 (1)				
	0.86 (1)	0.74 (1)	1.00 (1)				0.96 (1)	
			+ (1)					
4	12	16	12	25	6	9	8	5
105–108	105–116	117–132	137–148	149–173	174–179	193–201	202–209	210–214
Glu				Lys				

value given is for *S*-carboxymethylcysteine rather than for *S*-aminoethylcysteine.

represented incomplete tryptic cleavage at the Arg–Val bond at positions 77–78, as is shown by the independent isolation of TVt5′, Ile-Ser-Arg. The low yield of TVt5 is attributable to poor cleavage after the COOH-terminal aminoethylcysteine, which is frequent in our experience. The next peptide, TVt6, was also obtained in low yield (8%). The composite peptide TVt5–TVt6 (residues 78–103) was not isolated, probably because of its large size and hydrophobic character. The sequence in this region was assembled from a series of subpeptides of TVt5 and TVt6 and from the chymotryptic peptides of the whole protein.

A peptic digest of TVt5 yielded seven subpeptides (P1–P7) (Table VI). From the composition of TVt5-P1 and the Edman data on TVt5′, the sequence Ile-Ser-Arg-Val was deduced. The placement at the amino terminus depends on the composition of the series of overlapping subpeptides P1–P6, from which the sequence was deduced as Ile-Ser-Arg-Val (Glx,Ala)Glx-Asp-Val-Gly-Val. Because of the presence of the basic residue, aminoethylcysteine, TVt5-P7 was assumed to have the sequence Tyr-Tyr-Cys and was placed at the COOH terminus by homology to other κ chains in which this tripeptide sequence is almost invariant for positions 86–88


NH2-ASP-ILE-VAL-MET-THR-GLN-SER-PRO-LEU-SER-LEU-PRO-VAL-THR-PRO-GLY-GLU-PRO-ALA-SER-ILE-SER-CYS-ARG-SER-SER-GLN-SER-LEU-LEU-HIS-SER-ASP-GLY-PHE-
Sequenator
TVt1
TVt1-P1
TVt1-P2
TVt1-P3
TVt1-P4
TVt1-C1
TVt1-C2
TVt1-C3
Ct1
Ct4
Ct4'
Ct5
Ct3
CNBr-I
Ct2
ASP-TYR-LEU-ASN-TRP-TYR-LEU-GLN-LYS-PRO-GLY-GLN(SER,PRO,GLX)LEU-LEU-ILE(TYR)(ALA,LEU)SER-ASN-ARG-ALA-SER-GLY-VAL-PRO-ASP-ARG-
Sequenator
TVt3-TVt4
TVt3
Ct16
Ct8
Ct12
Ct7
Ct11
Ct14
Ct9
Ct11-Pap1
Ct11-Pap3
Ct15
Ct6
Ct11-Pap2
Ct11-Pr1
Ct11-Pr2
Ct10
Ct13
Legend
Edman
Dansyl
LAP
CPA
PHE-SER-GLY-SER-GLY-SER-GLY-THR-ASP-PHE-THR-LEU-LYS-ILE-SER-ARG-VAL-GLU-ALA-GLU-ASP-VAL-GLY-VAL-TYR-TYR-CYS-MET-GLX-ALA-LEU-
TVt5'
TVt4
TVt5
TVt4-P1
TVt4-P3
TVt5-P2
TVt5-P4
TVt5-P7
TVt4-P2
TVt5-P3
TVt5-P5
TVt4-Pap1
TVt4 Pap2
TVt4-Pap3
TVt4 Pap4
TVt5-P1
TVt5-P6
Ct 16
Ct17
Ct18
Ct20
Ct 14
Ct17-Pap1
Ct17-Pap2
Ct18-T1
Ct18-T2
Ct 15
Ct19
GLN-ALA-PRO-ILE-THR-PHE-GLY-GLN-GLY-THR-ARG-LEU-GLU-ILE-LYS-ARG(THR,VAL,ALA,ALA,PRO,SER,VAL,PHE)(ILE,PHE,PRO,PRO,SER,ASN,GLU,
TVt6
TVt7
Tk1
Tk2
TVt6-C1
TVt7-Tk1
Ct21
Ct22
Ct25
Ct23
Ct24
GLU,LEU)LYS-SER(GLY,THR,ALA,SER,VAL)VAL-CYS-LEU-LEU(ASN,ASN,PHE,TYR,PRO)ARG(GLU,ALA)LYS-VAL(GLN,TRP)LYS-VAL(ASP,ASN,ALA,LEU,
Tk3a
Tk3b
Tk4
Tk5
Ct26
Ct27
GLN,SER,GLY,ASN,SER,GLN,GLU,SER,VAL,THR,GLU,GLN,ASP,SER)LYS-ASP(SER,THR,TYR)(SER,LEU,SER,SER,THR,LEU)(THR,LEU,SER)LYS-ALA(ASP,
Tk6
Tk7
Ct28
Ct29

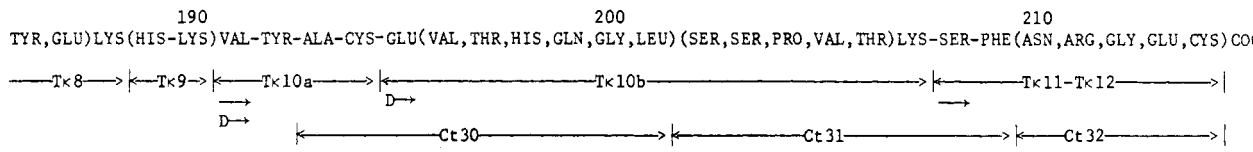


FIGURE 5: Amino acid sequence of the κII Bence Jones protein Tew as determined by sequenator analysis of the whole protein and sequence determination of the tryptic and chymotryptic peptides. The symbols are explained in the legend and the designations of the peptides in the text. Residues in the undetermined portion of the peptides of the constant region (Thr-109 through Cys-214) are aligned to accord with the sequence established for the constant region of human κ light chains (Putnam *et al.*, 1966; Putnam, 1969; Hilschmann, 1969). Broken arrows under His-30a and Asn-34 indicate some uncertainty in the identification of these residues in the sequenator analysis, but His-30a was established independently in Ct5 and Asn-34 in Ct9.

(Putnam, 1969). The order in this region was verified through the chymotryptic peptides Ct18–Ct20 (Figure 5). Since TVt5-P5 had acidic mobility in electrophoresis at pH 6.5 the Asx was aspartic acid.

Peptide TVt6. Two steps of Edman degradation gave the amino-terminal sequence Met-Glx (Figure 5). A chymotryptic digest of TVt6 yielded only one subpeptide that was pure enough for sequence analysis (TVt6-C1) Three steps of Edman degradation gave the sequence Gly-Glx-Gly for the amino terminus of TVt6-Cl. This subpeptide was basic on electrophoresis at pH 6.5 indicating the Glx was present as glutamine. From the specificity of trypsin, the arginine and thus this subpeptide must be COOH terminal in TVt6. This gives the partial sequence for TVt6 as Met-Glx(Ala,Leu,Glx,Ala,Pro,-Ile,Thr,Phe)Gly-Gln-Gly-Thr-Arg. Here the proof of the sequence depends on the chymotryptic peptides Ct20–Ct23; these also show that only one proline residue is present in TVt6.

Peptides TVt7, Tκ1, and TVt7–Tκ1. These three peptides conclude the V region of the Tew κ chain. Two steps of Edman degradation showed the amino-terminal sequence of TVt7 to be Leu-Glx (Table VII). The neutral mobility at pH 6.5 of TVt7 indicated the Glx to be glutamic acid. Lysine was placed as COOH terminal from the specificity of trypsin to give the sequence Leu-Glu-Ile-Lys. Free arginine (Tκ1) was isolated from the tryptic digest; this was placed in position 108 by considering TVt7–Tκ1, which had N-terminal leucine by the dansyl method and in which the Lys–Arg bond was resistant to trypsin. This overlap was verified by Ct24 and Ct25 (Figure 5). Because of its frequency in other human κ chains the arginine at position 108 was once thought to be the first residue of the C region and was designated Tκ1. However, glycine has been found at this position in one human κ chain, so the C region begins with Thr-109 (Gottlieb *et al.*, 1970).

Tryptic Peptides of the C Region (Tκ2–Tκ12). As summarized in Table VIII, all tryptic peptides expected for the constant region of the human κ chain were isolated from the Tew Bence Jones protein with the exception of the dipeptide His-Lys corresponding to positions 189–190. However, this was identified in the peptide map of Tew by its characteristic position and Pauly-positive reaction (Easley and Putnam, 1966). The stoichiometry of peptides Tκ2–Tκ12 agreed well with that expected from the sequence of the known tryptic peptides (Titani *et al.*, 1969). In the Tew protein peptides Tκ3 and Tκ10 were each subdivided into two peptides designated Tκ3a and Tκ3b and Tκ10a and Tκ10b, respectively, because the protein was aminoethylated, thus providing an additional bond susceptible to trypsin after aminoethylcysteine-134 and aminoethylcysteine-194. However, in accord with our previous experience with the Ou κ light chain (Putnam *et al.*, 1971), trypsin did not remove the COOH-terminal residue, aminoethylcysteine-214, nor did it cleave the Arg–Gly bond at 211–212 but instead gave the combined peptide Tκ11–Tκ12. The amino terminus of most of these peptides was determined by the Edman degradation or by the dansyl method to substantiate their identification as C-region peptides. The presence of valine at position 191 in Tκ10a accords with the serological classification of the Tew κ chain as an Inv (3) protein, Val-191 being associated with Inv (3) and Leu-191 with Inv (1,2) (Easley and Putnam, 1966).

Summary of the Tryptic Peptides. Of the 219 residues given in the sequence for the Tew κ chain in Figure 5, 182 are accounted for by the tryptic peptides summarized in Table VIII

TABLE III: Composition of Bence Jones Protein Tew from Amino Acid Analysis and Sequence Analysis and of Tissue Amyloid Protein.[a]

Amino Acid	Amino Acid Analysis	Nearest Integer	From Sequence	Amyloid Protein[f]
Lysine	11.4	11	11	11.9
Histidine	2.7	3	3	3.1
Arginine	7.7	8	8	8.3
Aspartic acid	18.1	18	18	19.7
Threonine	13.7	14	14	14.7
Serine	26.6	30[b]	31	25.2
Glutamic acid	25.5	26	25	28.5
Proline	11.8	12	13	13.4
Glycine	14.1	14	14	15.5
Alanine	13.0	13	13	14.4
Half-cystine	5[c]	5	5	*f*
Valine	15.6	16	16	16.4
Methionine	1.6	2	2	2.2
Isoleucine	6.5	7	7	7.2
Leucine	19.4	19	20	21.7
Tyrosine	7.4	8[d]	9	7.9
Phenylalanine	8.4	8	8	9.9
Tryptophan	(2)[e]	(2)[e]	2	*f*
Total			219	

[a] Values are given as moles per mole of protein. [b] Corrected for 10% destruction of serine during acid hydrolysis. [c] Determined as aminoethylcysteine. [d] Corrected for 5% destruction of tyrosine during acid hydrolysis. [e] Presence of two residues of tryptophan assumed from sequence analysis of whole protein and from qualitative staining reactions of tryptic and chymotryptic peptides. [f] Amino acid composition given by Terry *et al.* (1973) recalculated from mol/100 mol of amino acids to mol/220 mol of amino acids. Tryptophan and half-cystine were not determined in the amyloid protein.

TABLE IV: Sequence of Tryptic Peptide TVt1.

		1 24
Sequence:		Asp-Ile-Val-Met-Thr-Gln-Ser-Pro-Leu-Ser-Leu-Pro-Val-Thr-Pro-Gly-Glu(Pro,Ala)Ser-Ile-Ser-Cys-Arg[a]
Peptic subdigest:	P1	(Asp,Ile,Val,Met)
	P2	Thr-Gln-Ser-Pro-Leu-Ser-Leu
	P3	Pro-Val-Thr-Pro-Gly-Glu(Pro,Ala)Ser
	P4	Ile-Ser-Cys-Arg
Chymotryptic subdigest:	C1	(Asp,Ile,Val,Met)
	C2	(Thr,Glx,Ser,Pro,Leu)
	C3	Ser-Leu

HVE, pH 6.5: P1, P3, and C1, acidic; P2, neutral

[a] The order of the subpeptides was established from the chymotryptic peptides of the whole protein and by sequencer analysis of the whole protein.

TABLE V: Partial Sequence of Tryptic Peptide TVt4.

		62 74
Sequence:		Phe-Ser(Gly,Ser,Gly)(Ser,Gly)Thr-Asp-Phe-Thr-Leu-Lys
Peptic subdigest:	P1	(Phe,Ser,Gly,Ser,Gly,Ser,Gly,Thr,Asp,Phe)
	P2	(Phe,Thr,Leu)Lys
	P3	(Thr,Leu)Lys
Papain subdigest:	Pap1	(Phe,Ser,Gly,Ser,Gly)
	Pap2	(Ser,Gly)
	Pap3	Thr(Asp,Phe,Thr)
	Pap4	Leu-Lys
HVE, pH 6.5:		P1 and Pap3, acidic

TABLE VI: Partial Sequence of Tryptic Peptides TVt5 and TVt5′.

		75 88
Sequence:		Ile-Ser-Arg-Val(Glx,Ala)Glx-Asp-Val-Gly-Val-Tyr-Tyr-Cys
Peptic subdigest:	P1	(Ile,Ser,Arg,Val)
	P2	(Ile,Ser,Arg,Val,Glx,Ala)
	P3	(Ile,Ser,Arg,Val,Glx,Ala,Glx)
	P4	(Glx,Asx,Val,Gly)Val
	P5	(Asp,Val,Gly,Val)
	P6	Val(Gly,Val)
	P7	Tyr-Tyr-Cys
TVt5′		Ile-Ser-Arg
HVE, pH 6.5:		P5, acidic; dansyl gave isoleucine as the NH_2 terminus of TVt5

and more than 80 residues are positioned in sequence. This includes some 64 residues in the V region, the sequence determination of which was the main object of this study.

In the three human κ chains we have sequenced no tryptic peptides were identical in the V region except for the two representing the sequence Leu-Ile-Glu-Lys-Arg from positions 104 to 109; this short stretch at the end of the V region is identical in some human κ chains such as Tew, Ag (Titani *et al.*, 1969), and Ou (Putnam *et al.*, 1971) but differs in others. However, as indicated by the nomenclature Tκ2, Tκ3, etc., all tryptic peptides of the C region had identical counterparts in the three proteins. Nonetheless, because of the great similarity of Tew to other members of subgroup κII all the peptides of the V region of Tew could be aligned in order solely by homology to the known sequences of about ten other human κ chains. To get independent evidence for overlaps for the tryptic peptides and to fill out the missing portions, a chymotryptic digest was prepared and 33 chymotryptic (Ct) peptides were purified. Sequence analysis was done only on those of interest for areas of undetermined structure.

TABLE VII: Sequence of Tryptic Peptides TVt7, Tκ1, and TVt7-Tκ1.

	104 108
Sequence:	Leu-Glu-Ile-Lys-Arg
TVt7	Leu-Glu-Ile-Lys
Tκ1	Arg
TVt7-Tκ1	Leu(Glu,Ile,Lys,Arg)
HVE, pH 6.5:	D→ TVt7 neutral

Chymotryptic Peptides

Peptide Ct1. Two steps of Edman degradation gave the sequence Asx-Ile (Figure 5). Leucine aminopeptidase digestion released aspartic acid. These results plus the presence of methionine showed Ct1 to be at the amino terminus. As in TVtl, the low values for isoleucine and valine after acid hydrolysis for 22 hr reflect the stability of the Ile–Val bond. The low value for methionine is attributed to oxidative destruction.

Peptides Ct2, Ct3, and Ct4. These three peptides each begin with residue 10 and extend for different lengths as shown in Figure 5. They provide the overlaps for the peptic and chymotryptic subpeptides of TVtl (Table IV).[2] Ct4 extends the sequence of TVtl by three residues (Ser-Ser-Gln). However, these could not be differentiated in Ct4 because after carboxypeptidase A digestion only a peak at the position of serine was obtained on amino acid analysis. This indicated the absence of glutamic acid at the C-terminus of Ct4 and suggests the presence of glutamine, as was later confirmed by use of the sequencer on the whole protein.

Peptides Ct4' and Ct5. These two small peptides could not be identified in any of the tryptic peptides isolated. One step of Edman degradation gave the amino terminus of Ct5 as histidine; carboxypeptidase A digestion for 15 min released only phenylalanine giving the partial sequence of Ct5 as His(Ser, Asx,Gly)Phe. Placement was later achieved by use of the sequenator on the whole protein.

Peptides Ct6, Ct7, Ct8, and Ct9. The hexapeptide sequence represented by these four peptides (Table IX) caused great difficulty because of the multiple positions of chymotryptic cleavage and our inability to get good yields of the tryptophan-containing peptides. Another problem is that the tetrapeptide Ct8 (Leu,Asn,Trp,Tyr) could exist in three forms that would be hard to separate, *e.g.*, Tyr-Leu-Asn-Trp, Leu-Asn-Trp-Tyr, and Asn-Trp-Tyr-Leu. As a result, clear sequence data were obtained only on the dipeptide Ct9; this was shown to be Asx-Trp by Edman degradation. Since Ct7, Ct8, and Ct9 moved as neutral peptides in paper electrophoresis at pH 6.5,

[2] A noteworthy feature of the chymotryptic peptides is that Ct2, Ct12, and Ct19 require cleavage at the carboxyl group of 3 different isoleucine residues. Although not unprecedented, this is unusual. All of these peptides were obtained in yields too low for sequence analysis. This probably reflects the fact that although chymotrypsin has greatest specificity for cleavage of peptide bonds involving the carboxyl group of aromatic amino acids, the intrinsic specificity is rather broad; slow hydrolysis may occur at the carboxyl group of many different amino acids including isoleucine (Hill, 1965). Alternatively, the expected chymotryptic peptide Leu-Leu-Ile-Tyr may have been obtained in low yield because of its hydrophobicity, and the tyrosine may have been destroyed by oxidation during acid hydrolysis. Although the sequence of Ile-Tyr had to be assumed by analogy, the position of Ile-21 in Ct2 and of Ile-75 just before Ct19 are well established in other peptides.

TABLE VIII: Summary of the Tryptic Peptides Isolated from Bence Jones Protein Tew.

Peptide	Residue Position	No. of Residues	
TVt1	1–24	24	Asp-Ile-Val-Met-Thr-Gln-Ser-Pro-Leu-Ser-Leu-Pro-Val-Thr-Pro-Gly-Glu(Pro,Ala)Ser-Ile-Ser-Cys-Arg
TVt2	25–54	35	Not isolated; sequence determined on chymotryptic peptides and by use of the sequencer.
TVt3	55–61	7	Ala-Ser-Gly-Val-Pro-Asp-Arg
TVt4	62–74	13	Phe-Ser(Gly,Ser,Gly)(Ser,Gly)Thr-Asp-Phe-Thr-Leu-Lys
TVt5	75–88	14	Ile-Ser-Arg-Val(Glx,Ala)Glx-Asp-Val-Gly-Val-Tyr-Tyr-Cys
TVt6	89–103	15	Met-Glx(Ala,Leu,Glx,Ala,Pro,Ile,Thr,Phe)Gly-Gln-Gly-Thr-Arg
TVt7	104–107	4	Leu-Glu-Ile-Lys
Tκ1	108	1	Arg
Tκ2	109–126	18	(Thr,Val,Ala,Ala,Pro,Ser,Val,Phe,Ile,Phe,Pro,Pro,Ser,Asx,Glx,Glx,Leu)Lys
Tκ3a	127–134	8	Ser(Gly,Thr,Ala,Ser,Val,Val)Cys
Tκ3b	135–142	8	Leu-Leu(Asx,Asx,Phe,Tyr,Pro)Arg
Tκ4	143–145	3	(Glx,Ala)Lys
Tκ5	146–149	4	Val(Glx,Trp)Lys
Tκ6	150–169	20	Val(Asx,Asx,Ala,Leu,Glx,Ser,Gly,Asx,Ser,Glx,Glx,Ser,Val,Thr,Glx,Glx,Asx,Ser)Lys
Tκ7	170–183	14	Asp(Ser,Thr,Tyr,Ser,Leu,Ser,Ser,Thr,Leu,Thr,Leu,Ser)Lys
Tκ8	184–188	5	Ala(Asx,Tyr,Glx)Lys
Tκ9	189–190	2	(His-Lys). Not isolated, but identified in peptide map.
Tκ10a	191–194	4	Val(Tyr,Ala)Cys
Tκ10b	195–207	13	Glu(Val,Thr,His,Glx,Gly,Leu,Ser,Ser,Pro,Val,Thr)Lys
Tκ11-12	208–214	7	Ser(Phe,Asx,Arg,Gly,Glx,Cys)

TABLE IX: Alignment of Chymotryptic Peptides Ct6-Ct9.[a]

	31 36
Sequence:	Asx-Tyr-Leu-Asn-Trp-Tyr
Ct6	(Asx,Tyr,Leu,Asx,Trp)
Ct7	(Leu,Asn,Trp)
Ct8	(Leu,Asn,Trp,Tyr)
Ct9	Asn-Trp
HVE, pH 6.5:	Ct7, Ct8, and Ct9, neutral

[a] Alignment proposed for peptides; sequence in this area established by sequenator analysis.

the Asx is asparagine. The tentative sequence shown in the alignment in Table IX is not present in any of the tryptic peptides isolated but was established for positions 31–36 by use of the sequenator on the whole protein.

Peptides Ct10, Ct11, Ct12, and Ct13. This series of overlapping chymotryptic peptides (Table X and Figure 5) provided evidence for the sequence from positions 37 to 48 in the region covered by the missing tryptic peptide TVt2. Leucine aminopeptidase digestion of Ct10 released leucine, 1.0 mol/mol, and a residue in the position of serine on the amino acid analyzer 0.65 mol/mol. The latter peak must represent glutamine because the Edman data on Ct11 gave the sequence Leu-Glx. Papain digestion of Ct11 gave three subpeptides, Papl-Pap3, and pronase digestion gave two, Pr1 and Pr2. Four of these failed to react in the Edman degradation suggesting the presence of cyclized glutamine both at position 38 (confirming the conclusion on Ct10) and at position 42. Two steps of Edman degradation on Ct11-Pap2 gave Lys-Pro-Gly for this tripeptide. The carboxyl-terminal sequence for this region is Leu-Leu-Ile;[2] this conclusion is based on the release of two residues of leucine by carboxypeptidase A digestion of Ct13, the composition of Ct12, and the alignment of Ct10–Ct13 shown in Table X. From the cessation of action of carboxypeptidase A on Ct13 and by homology to other κ chains, the sequence from positions 43 to 45 is probably Ser-Pro-Gln, but this segment is left in parentheses because the evidence is indirect.

Although a chymotryptic peptide corresponding to the region from Gln-38 through Arg-54 was obtained, it could not be obtained in an amount sufficient for structural study. Thus, we have no direct evidence for the tyrosine residue in parentheses at position 49; this is placed there solely by homology because all other human and mouse κ chains sequenced have tyrosine in this position.

TABLE X: Partial Sequence of Chymotryptic Peptides Ct10, Ct11, Ct12, and Ct13.

	37 40 48
Sequence:	Leu-Gln-Lys-Pro-Gly-Gln(Ser,Pro,Glx)Leu-Leu-Ile
Ct10	Leu-Gln(Lys,Pro,Gly,Glx,Ser,Pro,Glx)
Ct11	Leu-Glx(Lys,Pro,Gly,Glx,Ser,Pro,Glx,Leu,Leu)
Ct11-Pap1	(Glx,Lys,Pro,Gly)[a]
Ct11-Pap2	Lys-Pro-Gly
Ct11-Pap3	(Glx,Ser,Pro,Glx,Leu)[a]
Ct11-Pr1	(Glx,Lys,Pro,Gly)[a]
Ct11-Pr2	(Glx,Ser,Pro,Glx)[a]
Ct12	(Leu,Leu,Ile)
Ct13	(Lys,Pro,Gly,Glx,Ser,Pro,Glx)Leu-Leu

[a] Edman degradation indicated a blocked end-group suggestive of cyclized glutamine.

TABLE XI: Sequence of Chymotryptic Peptide Ct17.

		63 71
Sequence:		Ser-Gly-Ser-Gly-Ser-Gly-Thr-Asx-Phe
Papain subdigest:	Pap1	(Ser,Gly,Ser,Gly)
	Pap2	Ser-Gly-Thr(Asx,Phe)

Peptides Ct14, Ct15, and Ct16. These peptides serve to complete the missing sequence of TVt3 and provide the overlap to TVt3 and TVt4 (Figure 5). The three peptides all terminate with Phe-62. Ct14 contains an alanine and a leucine residue not present in the others, so they must be at the amino terminus as shown by the alignment in Figure 5. Insufficient Ct14 was available for structural study so these residues are placed in parentheses at positions 50 and 51; however, the order is probably Ala-Leu because of the chymotryptic action that gave Ct15. Edman degradation of Ct15 gave serine as the amino terminus. Aspartic acid was not released by leucine aminopeptidase, but the serine peak in the analyzer was greater than unity indicating that aspargine was released. Since TVt3 ought to be preceded by a basic residue, the arginine was placed at position 54 giving the amino-terminal sequence of Ser-Asn-Arg for Ct15. This accords with the relationship of Ct15 and Ct16 and the proof of overlap given by TVt3–TVt4.

Peptide Ct17. Complete sequence analysis was done on Ct17 because of the troublesome repetitious Ser-Gly sequence shown by TVt4. Four rounds of the Edman degradation established the amino-terminal sequence of Ct17 as Ser-Gly-Ser-Gly (Table XI). Digestion of Ct17 with papain yielded two subpeptides. Ct17-Papl (Ser,Gly,Ser,Gly) was identical with the four residues shown in the amino terminus by Edman degradation. Ct17-Pap2 represented the remainder of Ct17. Three steps of Edman degradation gave the amino terminus as Ser-Gly-Thr. The last two residues were arranged as Asx-Phe because carboxypeptidase A digestion released only phenylalanine. The Asx had been identified as aspartic acid in the tryptic peptide TVt4.

A chymotryptic peptide corresponding to residues 72–74 (Thr-Leu-Lys) was not recovered.

Peptides Ct18 and Ct19. The sequence from 75 to 84 was determined on the tryptic peptides TVt5′ and TVt5 except for the arrangement of (Glx,Ala) in positions 79–80. Because of the presence of arginine a tryptic digestion of Ct18 was made, which yielded two subpeptides, Ct18-Tl and Ct18-T2 (Table XII). The former (Ile,Ser)Arg is identical with the tryptic peptide TVt5′ isolated from the digest of the whole protein. Two steps of Edman degradation on Ct18-T2 showed the amino terminus to be Val-Glx; this arranges the (Glx,Ala)

TABLE XII: Partial Sequence of Chymotryptic Peptides Ct18 and Ct19.

	75 80 86
	Ile-Ser-Arg-Val-Glu(Ala,Glu,Asp,Val,Gly,Val,Tyr)
Ct18	(Ile,Ser,Arg,Val,Glx,Ala,Glx,Asx,Val,Gly,Val,Tyr)
Ct18-T1	(Ile,Ser)Arg
Ct18-T2	Val-Glu(Ala,Glu,Asp,Val,Gly,Val,Tyr)
Ct19	(Ser,Arg,Val,Glx,Ala,Glx,Asx,Val,Gly,Val,Tyr)

TABLE XIII: Summary of Sequence Data on the Chymotryptic Peptides of Protein Tew.

Peptide	Residue Position	No. of Residues	Sequence
Ct1	1–9	9	Asp-Ile(Val,Met,Thr,Glx,Ser,Pro,Leu)
Ct5	30a–30e	5	His(Ser,Asx,Gly)Phe
Ct9	34–35	2	Asn-Trp
Ct11	37–47	11	Leu-Glx-Lys-Pro-Gly-Gln(Ser,Pro,Glx)Leu-Leu
Ct15	52–62	11	Ser-Asn(Arg,Ala,Ser,Gly,Val,Pro,Asx,Arg,Phe)
Ct17	63–71	9	Ser-Gly-Ser-Gly-Ser-Gly-Thr-Asp-Phe
Ct18	75–86	12	(Ile,Ser)Arg-Val-Glu(Ala,Glu,Asp,Val,Gly,Val,Tyr)
Ct21	93–98	6	Gln-Ala-Pro-Ile-Thr-Phe

residues in question in TVt5 as Glx-Ala. To determine the form of the Glx residue, leucine aminopeptidase digestion was done on Ct18-T2. The enzyme completely digested the subpeptide releasing two residues of glutamic acid and one of aspartic acid, as well as the nonacidic residues. These results indicate the sequence in question is Glu-Ala and that the other Glx and the Asx are also present in the acidic form. Except for the absence of isoleucine[2] Ct19 is identical in composition with Ct18; this confirms that isoleucine must be amino-terminal in Ct18 and that the sequence of Ct18-T1 is Ile-Ser-Arg.

Peptide Ct20. This peptide containing *S*-carboxymethylcysteine and methionine provided the overlap between TVt5, which had COOH-terminal aminoethylcysteine, and TVt6. Because of the specificity of chymotrypsin, leucine was assumed to be COOH terminal in Ct20. Insufficient peptide was available for further study.

Peptide Ct21. This peptide was completely sequenced because it was from an area in the tryptic peptide TVt6 that had not been determined (Figure 5). Three steps of the Edman degradation established the amino terminus as Glx-Ala-Pro. Leucine aminopeptidase digestion of Ct21 released glutamine, which was identifiable on the amino acid analyzer because of the absence of serine in this peptide. Carboxypeptidase A digestion of Ct21 for 6 hr released phenylalanine and threonine, 1.0 and 0.38 mol per mol, respectively to show the carboxyl-terminal sequence as Thr-Phe. This places isoleucine as the fourth residue in Ct21 to give the sequence Gln-Ala-Pro-Ile-Thr-Phe.

Peptides Ct22 and Ct23. Peptide Ct22 extends from position 99 to 104 and Ct23 from 100 to 104. Thus, both give an overlap of TVt6 and TVt7 and by difference confirm the glycine at position 99.

Peptides Ct24 and Ct25. Both peptides begin at position 105, Ct24 extending to position 108 and Ct25 to position 116. The two peptides provide an overlap of TVt7 and Tκ1, and Ct25 extends this into the constant region beginning with Tκ2.

Peptide Ct26. This peptide contains parts of Tκ2 and Tκ3a and provides proof of their overlap. A chymotryptic peptide giving the overlap of Tκ3a and Tκ3b was not isolated but the sequence here, Val-Cys-Leu-Leu, is invariant in human κ chains.

Peptide Ct27. The overlap of Tκ3b, Tκ4, and Tκ5 is provided by this peptide.

Peptide Ct28. This 25-residue peptide covering residues 149–173 was the largest peptide obtained from protein Tew. It overlaps Tκ5, Tκ6, and Tκ7. The same chymotryptic peptide was isolated from the human κ chains Ou (Putnam *et al.*, 1971) and Ag (Whitley *et al.*, 1969). It was the only chymotryptic peptide isolated from this area in all three κ chains although a number of chymotryptic-sensitive residues such as a leucine and several glutamines are present. This region may be in a protected area due to tertiary structure of the protein.

Peptide Ct29. This is wholly contained in Tκ7. Chymotryptic peptides for the following region (residues 180–192) were not isolated.

Peptides Ct30, Ct31, and Ct32. Together these peptides complete the carboxyl terminus of this κ chain. Ct30 gives the overlap for Tκ10a and Tκ10b and Ct31 for Tκ10b and Tκ11. Ct32, which is in Tκ11–Tκ12, is important because a good recovery was obtained (0.96 mol/mol) of the aminoethylcysteine residue which was missing in Tκ11–Tκ12.

Summary of the Chymotryptic Peptides. Since the main objective in isolation of the chymotryptic peptides was to secure overlaps for the tryptic peptides and to fill in the missing areas in the V region, relatively little sequence work was done on them. Except for peptides for which only end group data were obtained, the results are summarized in Table XIII.

In the V regions of the three human κ chains we have sequenced no chymotryptic peptides are identical except for the tripeptide Ct7 in Tew and Ca3 in Ag (Whitley *et al.*, 1969); this represents the highly conserved Leu-Asn-Trp sequence seen in many human κ chains. However, identical chymotryptic peptides covering much of the C region were isolated from both Tew (Ct) and Ag (Ca), *e.g.*, Ct25(Tew) = Ca18 (Ag) residues 105–116; Ct27 = Ca23 (137–148); Ct28 = Ca24 (149–173); Ct29 = Ca25 (174–179); Ct30 = Ca32 (193–201); Ct31 = Ca34 (202–209); Ct32 = Ca35 (210–214).

Analysis of the Tew κ Chain with the Protein Sequencer

After completion of the work described above on the tryptic and chymotryptic peptides, there was insufficient Tew protein remaining to repeat the digests; yet, a number of positions in the sequence of the V region were uncertain and some overlaps were not established. The problem was most serious for the missing tryptic peptide TVt2 where several gaps existed in the sequence and the overlaps of the chymotryptic peptides were uncertain. Furthermore, because the Tew κ chain closely resembled the κII subgroup proteins Mil and Cum in sequence, it was important to ascertain whether, like them, Tew had extra residues after position 30 (*e.g.*, 30a, 30b, etc.) or whether the extra residues were absent in Tew as they are in all κI and κIII proteins thus far sequenced. Because this characteristic is important as a criterion for the classification of VκII subgroup proteins, we deferred publication of the Tew κ sequence until the question could be settled.

After we obtained the Beckman Model 890 protein sequenator, three attempts were made to sequence the Tew κ chain by the automatic method. At this point the chymo-

TABLE XIV: Acid and Amide Assignments of the Aspartic and Glutamic Acid Residues from the Variable Region of Protein Tew.[a]

Residue No.	Assignment	Method of Determination
1	Asp	Hve, TVt1-P1, TVt1-C1; LAP, Ct1, GC
6	Gln	Hve, TVt1-P2; LAP, TVt1-P2; GC
17	Glu	Hve, TVt1-P3; GC
27	Gln	CPA, Ct4; GC
30c	Asp	GC
31	Asp	GC
34	Asn	Hve, Ct7, Ct8, Ct9; GC (probable)
38	Gln	LAP, Ct10; Edman degradation indicated PCA for Ct11-Pap1 and Ct11-Pr1
42	Gln	Edman degradation indicated PCA for Ct11-Pap3 and Ct11-Pr2
45	Glx	
53	Asn	LAP, Ct15
60	Asp	Hve, TVt3
70	Asp	Hve, TVt4-P1 and TVt4-Pap3
79	Glu	LAP, Ct18-T2
81	Glu	LAP, Ct18-T2
82	Asp	LAP, Ct18-T2; hve, TVt5-P5
90	Glx	
93	Gln	LAP, Ct21
100	Gln	Hve, TVt6-C1
105	Glu	Hve, TVt7

[a] Abbreviations used are: hve, high-voltage electrophoresis at pH 6.5; CPA, carboxypeptidase A; LAP, leucine aminopeptidase; GC, gas chromatography of Pth derivative from sequencer; PCA, pyrrolidonecarboxylic acid.

tryptic peptides identified in Figure 5 as Ct4′ through Ct13 were not yet aligned, and the objective was to establish the first 40 steps in the sequence.

The fully reduced and carboxymethylated κ chain (10–15 mg) was submitted three times to automatic sequence analysis with the protein program. Gas and thin-layer chromatography of the phenylthiohydantoin derivatives and amino acid analysis of the recovered amino acids were used to identify the residues. The first two runs failed to give results after the 27th cycle. In accord with the carboxypeptidase data on Ct4, Glx-27 was shown to be glutamine by gas chromatography of the silylated derivative. It appeared that cyclization of Gln-27 blocked further Edman degradation by the automatic method. In a third attempt with 12 mg of reduced-carboxymethylated protein the acid cleavage time was shortened to 60 sec and successful degradation was achieved for 42 steps when again a glutamine (Gln-38) blocked further reaction.

As shown in Figure 5, analysis with the sequenator gave the overlap of TVt1 with TVt2, and permitted the ordering of Ct4–Ct9 and extension of the continuous sequence through Leu-37 including the extra residues 30a–30e. There is only one overlap (Leu-37) between Ct11 (and Ct10) and the sequence obtained with the sequenator, and there is no overlap between Ct10-Ct11-Ct12-Ct13 and the segments C terminal to this set of chymotryptic peptides. Therefore, the placement of the segment Gln-38 to Tyr-49 must rely largely on analogy. Most of the residues were identified directly by gas chromatography of the Pth derivatives. Gas chromatography of the silylated derivatives gave identification of Gln-6, Gln-27, Asp-30c, and Asp-31. The leucine residues were verified by amino acid analysis from position 29 on, as was Arg-24, Cm-cysteine-23, and His-30a. Although a number of the serines were ambiguous in the sequenator analysis, these positions were clear by reference to the peptides that had been sequenced manually.

An unsuccessful attempt was made to cleave the κ chain after Trp-35 by use of *N*-succinyl bromide in order to obtain a fragment beginning with Tyr-35 that was suitable for automatic sequence analysis. Because only a few mg of protein remained and were needed as a serological standard, it was not possible to complete the sequence for the six residues that are in parentheses from positions 42–52. However, they can all be ordered by homology to other κII chains.

Acid and Amide Assignments for Dicarboxylic Acid Residues

Assignment of the acid or amide form has been made for 18 of the 20 dicarboxylic acid residues in the V region of the Tew κ chain. The evidence for each residue was given in the text, and a summary is given in Table XIV. The undetermined residues are Glx-45 and Glx-90; by homology to other κ chains, both of these are probably glutamine. Gln-90 is invariant in all other human κ chains sequenced, and Gln-45 is present in the other κII chains Cum and Mil (Hilschmann, 1969).

Discussion

Although about a dozen human κ light chains have been sequenced since we reported the first complete sequence of any κ chain (Putnam *et al.*, 1966), the sequence of the Tew κ Bence Jones protein is of interest for two reasons. (1) It has unusually close homology to the Cum and Mil κ chains (Hilschmann, 1969) which are of subgroup κII, and thus provides new evidence for subgroup classification. (2) It was excreted by a patient who died of primary amyloidosis and it is apparently identical with the tissue amyloid protein of that patient.

The composition, partial sequence data, and alignment of the tryptic and chymotryptic peptides of the constant region (residues 109–214) of the Tew Bence Jones protein (Figure 5) correspond exactly to the sequence established for human κ light chains of the same allotype (Putnam *et al.*, 1966; Putnam,

	23								30	a	b	c	d	e	f	31								35
Tew	CYS	ARG	SER	SER	GLN	SER	LEU	LEU	HIS	SER	—	—	ASP	GLY		PHE	ASP	TYR	LEU	ASN	TRP	TYR	LEU	GLN
Cum	CYS	ARG	SER	SER	GLN	SER	LEU	LEU	ASP	SER	GLY		ASP	GLY		ASN	THR	TYR	LEU	ASN	TRP	TYR	LEU	GLN
Mil	CYS	ARG	SER	SER	GLN	ASN	LEU	LEU	GLX	SER	—	—	ASX	GLY		—(ASX)		TYR	LEU	ASP	TRP	TYR	LEU	GLX
Ag	CYS	GLN	ALA	SER	GLN	ASP	ILE	ASN	—	—	—	—	—	—	—	—	HIS	TYR	LEU	ASN	TRP	TYR	GLN	GLN

FIGURE 6: Comparison of the amino acid sequence of human light chains of the κII subgroup (Tew, Cum, and Mil) with the κI protein Ag used as the reference standard for the numbering system of human κ chains. The sources of sequence data are: Cum (Hilschmann, 1969), Mil (Dreyer *et al.*, 1967), Ag (Putnam *et al.*, 1966).

	54						60										70	
κII TEW	ARG	ALA	SER	GLY	VAL	PRO	ASP	ARG	PHE	SER	GLY	SER	GLY	SER	GLY	THR	ASP	PHE
κII MIL	ARG	ALA	SER	GLY	VAL	PRO	ASN	ARG	PHE	SER	GLY	SER	GLY	SER	GLY	THR	ASX	PHE
κII CUM	ARG	ALA	SER	GLY	VAL	PRO	ASP	ARG	PHE	SER	GLY	SER	GLY	SER	GLY	THR	ASP	PHE
κI ROY	LEU	GLU	ALA	GLY	VAL	PRO	SER	ARG	PHE	SER	GLY	THR	GLY	SER	GLY	THR	ASP	PHE
κI AG	LEU	GLU	THR	GLY	VAL	PRO	SER	ARG	PHE	SER	GLY	SER	GLY	PHE	GLY	THR	ASP	PHE
κI EU	LEU	GLX	SER	GLY	VAL	PRO	SER	ARG	PHE	ILE	GLY	SER	GLY	SER	GLY	THR	GLX	PHE

	72							80									
κII TEW	THR	LEU	LYS	ILE	SER	ARG	VAL	GLU	ALA	GLU	ASP	VAL	GLY	VAL	TYR	TYR	CYS
κII MIL	THR	LEU	LYS	ILE	SER	ARG	VAL	GLX	ALA	GLX	ASX	VAL	GLY	VAL	TYR	TYR	CYS
κII CUM	THR	LEU	LYS	ILE	SER	ARG	VAL	GLU	ALA	GLX	ASX	VAL	GLY	VAL	TYR	TYR	CYS
κI ROY	THR	PHE	THR	ILE	SER	SER	LEU	GLN	PRO	GLU	ASP	ILE	ALA	THR	TYR	TYR	CYS
κI AG	THR	PHE	THR	ILE	SER	GLY	LEU	GLN	PRO	GLU	ASP	ILE	ALA	THR	TYR	TYR	CYS
κI EU	THR	LEU	THR	ILE	SER	SER	LEU	GLX	PRO	ASX	ASX	PHE	ALA	THR	TYR	TYR	CYS

FIGURE 7: Comparison of the amino acid sequence of six human κ light chains from position 54 through 88. Three proteins are of the κII subgroup: Tew, Mil (Dreyer *et al.*, 1967), and Cum (Hilschmann, 1969), and three are of the κI subgroup: Roy (Hilschmann, 1969), Cum (Hilschmann, 1969), and Eu (Gottlieb *et al.*, 1970). The three κII proteins probably have an identical sequence (Met-Gln) at positions 89–90 since Hilschmann has reversed his earlier sequence here.

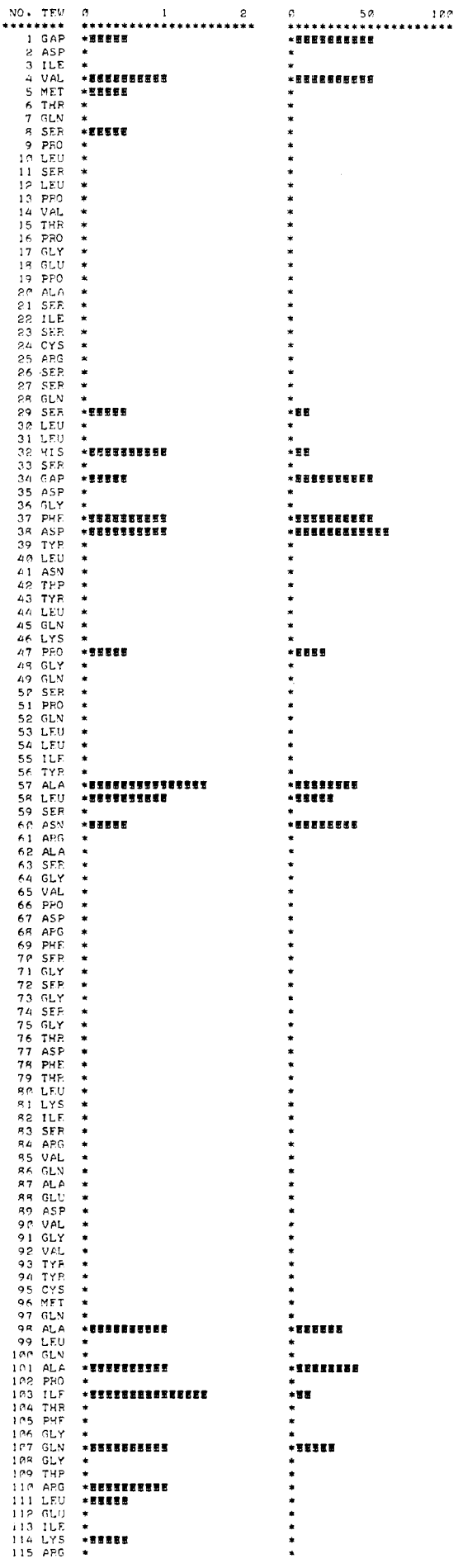
1969; Hilschmann, 1969). Hence, further discussion will be restricted to the sequence of the variable region (residues 1–108), which was the primary object of this study.

Bence Jones proteins Cum (Hilschmann, 1967), Mil (also designated HB3) (Dreyer *et al.*, 1967), and Tew are the only human light chains of the κII subgroup that have been completely or nearly completely sequenced. Although the workers who sequenced these proteins agreed on the numbering of the subgroups, the numbering was unfortunately reversed by an *ad hoc* group [*Bull. W.H.O. 41*, 975 (1969)]. In the latter, subgroup κII was designated $V_{\kappa III}$, and *vice versa* κIII was called $V_{\kappa II}$. Referring to the *WHO* bulletin, the Subcommittee for Human Immunoglobulins of the IUIS Nomenclature Committee has pointed out that "The nomenclature earlier proposed for subgroups needs revision" [*Biochemistry 11*, 3311 (1972)]. Hence, the original designation of subgroup κII (Putnam, 1969), equivalent to subgroup II of Milstein (1969), has been used in this article.

Three basic sequences (Milstein, 1969) or prototype sequences (*WHO* bulletin) for the first 20 residues of the amino terminus have been proposed for the classification of human κ chains into three subgroups. By chance the amino-terminal sequence of the Tew κ chain is identical to that postulated for the prototype sequence of the κII subgroup (Milstein, 1969) and the VκIII subgroup (*WHO* bulletin). The other proteins in this subgroup differ from the prototype sequence by one residue (Mil) and by three (Cum).

A second characteristic of the κII subgroup is the presence of additional residues relative to the accepted numbering system based on the sequence of the κI protein Ag (Titani *et al.*, 1969). In this system the invariant half-cysteine at position 23 and the invariant tryptophan at position 35 are taken as reference points. When this is done (Figure 6), the Tew κ chain has 5 extra residues after position 30, Cum has 6, and Mil has 4. Eight other human κ chains that have been fully sequenced have no extra residues after position 30 and fit in either subgroup κI or κIII according to their basic sequences. However, some human λ chains and some mouse κ chains have extra residues in the same area (Putnam, 1969).

The unusual degree of homology of the V region of three κII chains Tew, Mil, and Cum is illustrated in Figure 7. These three proteins are identical for 37 consecutive residues from position 54 through 90. This is the greatest stretch of identity yet identified for the V region of either human κ or λ light


NO. TEW 0 1 2 0 50 100
1 GAP
2 ASP
3 ILE
4 VAL
5 MET
6 THR
7 GLN
8 SER
9 PRO
10 LEU
11 SER
12 LEU
13 PRO
14 VAL
15 THR
16 PRO
17 GLY
18 GLU
19 PRO
20 ALA
21 SER
22 ILE
23 SER
24 CYS
25 ARG
26 SER
27 SER
28 GLN
29 SER
30 LEU
31 LEU
32 HIS
33 SER
34 GAP
35 ASP
36 GLY
37 PHE
38 ASP
39 TYR
40 LEU
41 ASN
42 TRP
43 TYR
44 LEU
45 GLN
46 LYS
47 PRO
48 GLY
49 GLN
50 SER
51 PRO
52 GLN
53 LEU
54 LEU
55 ILE
56 TYR
57 ALA
58 LEU
59 SER
60 ASN
61 ARG
62 ALA
63 SER
64 GLY
65 VAL
66 PRO
67 ASP
68 ARG
69 PHE
70 SER
71 GLY
72 SER
73 GLY
74 SER
75 GLY
76 THR
77 ASP
78 PHE
79 THR
80 LEU
81 LYS
82 ILE
83 SER
84 ARG
85 VAL
86 GLN
87 ALA
88 GLU
89 ASP
90 VAL
91 GLY
92 VAL
93 TYR
94 TYR
95 CYS
96 MET
97 GLN
98 ALA
99 LEU
100 GLN
101 ALA
102 PRO
103 ILE
104 THR
105 PHE
106 GLY
107 GLN
108 GLY
109 THR
110 ARG
111 LEU
112 GLU
113 ILE
114 LYS
115 ARG

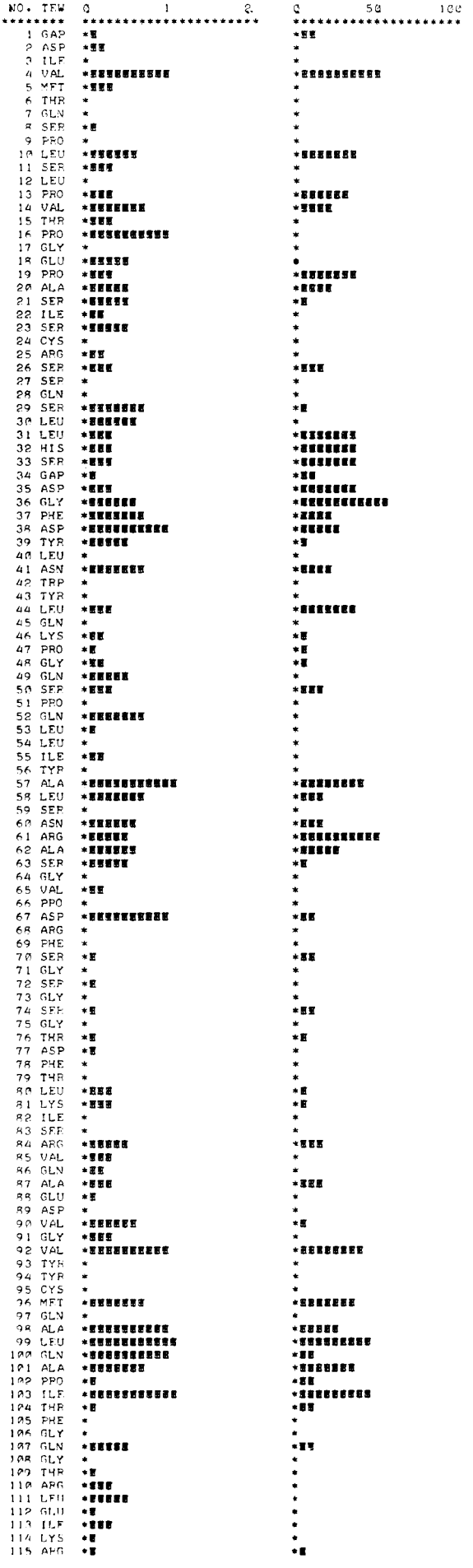

NO. TEW 0 1 2 0 50 100
1 GAP
2 ASP
3 ILE
4 VAL
5 MET
6 THR
7 GLN
8 SER
9 PRO
10 LEU
11 SER
12 LEU
13 PRO
14 VAL
15 THR
16 PRO
17 GLY
18 GLU
19 PRO
20 ALA
21 SER
22 ILE
23 SER
24 CYS
25 ARG
26 SER
27 SER
28 GLN
29 SER
30 LEU
31 LEU
32 HIS
33 SER
34 GAP
35 ASP
36 GLY
37 PHE
38 ASP
39 TYR
40 LEU
41 ASN
42 TRP
43 TYR
44 LEU
45 GLN
46 LYS
47 PRO
48 GLY
49 GLN
50 SER
51 PRO
52 GLN
53 LEU
54 LEU
55 ILE
56 TYR
57 ALA
58 LEU
59 SER
60 ASN
61 ARG
62 ALA
63 SER
64 GLY
65 VAL
66 PRO
67 ASP
68 ARG
69 PHE
70 SER
71 GLY
72 SER
73 GLY
74 SER
75 GLY
76 THR
77 ASP
78 PHE
79 THR
80 LEU
81 LYS
82 ILE
83 SER
84 ARG
85 VAL
86 GLN
87 ALA
88 GLU
89 ASP
90 VAL
91 GLY
92 VAL
93 TYR
94 TYR
95 CYS
96 MET
97 GLN
98 ALA
99 LEU
100 GLN
101 ALA
102 PRO
103 ILE
104 THR
105 PHE
106 GLY
107 GLN
108 GLY
109 THR
110 ARG
111 LEU
112 GLU
113 ILE
114 LYS
115 ARG

FIGURE 8 (opposite page): Computer comparison of the amino acid sequence of human κ light chains as a minimum position variation histogram. (A, left) Protein Tew compared to two other proteins of the κII subgroup (Mil and Cum). (B, right) Protein Tew compared to eight other human κ light chains—two of subgroup κII, four of subgroup κI, and two of subgroup κIII. In parts A and B the ordinate gives the amino acid sequence of protein Tew with gaps inserted to achieve maximum homology to other κ chains. The computer handles the gap as an amino acid and numbers the sequence serially and independent of the biochemical numbering system. In each figure the first histogram on the right gives the minimum nucleotide mutation necessary to convert one amino acid to another (scale from 0 to 2), and the second histogram gives the minimum Epstein value (scale 0 to 100). For further explanation of the method and interpretation see the Discussion.

chains. In contrast, proteins of the κI subgroup differ from κII at many positions shown in Figure 7 and to a lesser extent differ from each other in this region.

Although subgroup relationships can be demonstrated as above through similarity in the prototype amino-terminal sequence, through the presence or absence of extra residues after position 30, and through their overall homology, there are no clear quantitative criteria for defining the three κ-chain subgroups. In collaboration with Jeffrey N. Davidson of the Biology Department of Harvard University, we have undertaken a computer analysis of κ and λ light chains to ascertain the validity of the subgroup concept and to devise criteria for the assignment of light chains to subgroups.

By a computer program to be described elsewhere by Mr. Davidson, histograms are produced that compare light chains on the basis of similarities and differences in amino acid sequence. The homologous V region sequences are prealigned side by side before being entered into a computer. Since the sequences are of different length, gaps are inserted manually to achieve maximum homology. The computer handles a gap as an amino acid and numbers the sequence serially and independent of the biochemical numbering system. The computer compares corresponding amino acids at each position first in terms of minimum nucleotide mutations necessary to convert one amino acid to another (Fitch and Margoliash, 1967) and then again with respect to differences in amino acid polarity and size according to a matrix prepared by Epstein (1967). To minimize possible side effects the amino acids at each position are compared among themselves to give maximum fit, that is, each amino acid is compared to all others at that position to give the minimum nucleotide difference for the set and the minimum Epstein value. This is called the minimum position variation.

Computer-produced histograms are shown in Figure 8A for the κII proteins Tew, Mil, and Cum and in Figure 8B for these proteins plus six other human κ chains. Four of the latter are of the κI type, *i.e.*, Roy (Hilschmann, 1969), Ag (Titani *et al.*, 1969), Eu (Gottlieb *et al.*, 1970), and Ou (Putnam *et al.*, 1971); two are of the κIII type, *i.e.*, Ti (Hilschmann, 1969) and B6 (Milstein, 1969). From Figure 8 it is evident that Tew, Mil, and Cum are much more related to each other in amino acid sequence than they are to the other κ chains. In terms of minimum nucleotide mutations Tew differs from Cum and Mil by an average of only 0.2 base at each position, whereas the other human κ chains differ from Tew by an average from 0.4 to 0.6 base at each position. By the same criterion the V regions of the mouse κ chains M70 and M40 (Dreyer *et al.*, 1967) differ from the V region of Tew by 0.5 and 0.7 minimum nucleotide mutation at each position. Thus, the computer comparison gives a quantitative index of the homology of the V region of these light chains. By this index the V region of the mouse κ light chain M70 is as homologous to the V region of the three human κII light chains as are the V regions of most human κI or κIII chains. This suggests that the genes for Vκ subgroups do not arise by somatic mutation within individuals of the species but rather that they diverged early in evolution.

Two types of amyloid protein have recently been identified in the tissues of patients with amyloidosis. One of these proteins, amyloid protein A, seems characteristic of chronic or secondary amyloidosis and is unrelated to immunoglobulins in structure. Human amyloid protein A contains 77 residues in a single polypeptide chain devoid of disulfide bonds; its complete sequence has been determined by Levin *et al.* (1973) and the partial sequence by Ein *et al.* (1972). The complete sequence of monkey amyloid protein A, which is closely homologous, was earlier reported by Hermodson *et al.* (1972b).

A second type of amyloid protein that is related serologically and in primary structure to immunoglobulins has been isolated from patients with amyloidosis associated with plasma cell dyscrasias (Glenner *et al.*, 1972; Terry *et al.*, 1973). Amyloid fibril proteins isolated from such patients resemble κ or λ light chains in their amino-terminal sequence, and like light chains they differ in primary structure from one patient to another. However, the preponderance of amyloid proteins from patients with plasma cell dyscrasias appear to be light-chain fragments of the V region.

The unique finding in this work in collaboration with Drs. Osserman, Terry, and Glenner (Terry *et al.*, 1973) is that the predominant protein present in the amyloid deposits of patient Tew was an intact κ light chain that appears to be identical to the urinary Bence Jones protein studied by us. Using the protein sequenator, Dr. William Terry of the National Cancer Institute independently determined the sequence of the first 27 residues of a sample of the Tew Bence Jones protein which we provided. For both the Bence Jones protein and the amyloid protein the sequence is identical with that reported by us in Figure 5 for Asp-1 through Gln-27. The amino acid composition of the amyloid protein is closely similar to that of the Tew Bence Jones protein (Table III). Peptide maps of the two proteins are indistinguishable, indicating that no major insertions or deletions are present in the portion of the amyloid protein not directly sequenced. A tripeptide with a composition corresponding to the expected COOH-terminal tripeptide Gly-Glu-Cys was isolated from tryptic digests of the Bence Jones protein and the amyloid protein by Terry *et al.* (1973). The apparent identity of the urinary Bence Jones protein and the predominant tissue amyloid protein of this patient indicates the importance of immunologic phenomena in the pathogenesis of primary amyloidosis. Further study will be required to ascertain whether the deposition of the light chain as amyloid protein is the result of a specific immunologic reaction of the autoimmune type or whether deposition occurs nonspecifically because of the physicochemical properties of the protein and its abundant production.

Acknowledgment

We thank Dr. Elliott F. Osserman for supplying the Tew Bence Jones protein and for discussion of this case, and we thank Dr. William D. Terry for exchanging data on the Bence Jones protein and the amyloid protein from this patient.

References

Bernier, G. M., and Putnam, F. W. (1964), *Biochim. Biophys. Acta 86*, 295.
Dreyer, W. J., Gray, W. R., and Hood, L. (1967), *Cold Spring Harbor Symp. Quant. Biol. 32*, 353.
Easley, C. W., and Putnam, F. W. (1966), *J. Biol. Chem. 241*, 3671.
Ein, D., Kimura, S., Terry, W. D., Magnotta, J., and Glenner, J. D. (1972), *J. Biol. Chem. 247*, 5653.
Epstein, C. J. (1967), *Nature (London) 215*, 355.
Fitch, W. M., and Margoliash, E. (1967), *Science 155*, 279.
Gottlieb, P. D., Cunningham, B. A., Rutishauser, U., and Edelman, G. M. (1970), *Biochemistry 9*, 3155.
Gray, W. R. (1967), *Methods Enzymol. 2*, 469.
Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972a), *Biochemistry 11*, 4493.
Hermodson, M. A., Kuhn, R. W., Walsh, K. A., Neurath, H., Eriksen, N., and Benditt, E. P. (1972b), *Biochemistry 11*, 2934.
Hill, R. L. (1965), *Advan. Protein Chem. 20*, 37.
Hilschmann, N. (1967), *Hoppe Seyler's Z. Physiol. Chem. 348*, 1718.
Hilschmann, N. (1969), *Naturwissenschaften 56*, 195.
Köhler, H., Shimizu, A., Paul, C., and Putnam, F. W. (1970), *Science 169*, 56.
Levin, M., Franklin, E. C., Frangione, B., and Pras, M. (1973), *Biochemistry 12*, (in press).
Milstein, C. (1969), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 2*, 301.
Osserman, E. F., Takatsuki, K., and Talal, N. (1964), *Seminars Hematol. 1*, 3.
Pisano, J. J., and Bronzent, T. J. (1969), *J. Biol. Chem. 244*, 5597.
Putnam, F. W. (1969), *Science 163*, 633.
Putnam, F. W., Shimizu, A., Paul, C., Shinoda, T., and Köhler, H. (1971), *Ann. N. Y. Acad. Sci. 190*, 83.
Putnam, F. W., Titani, K., and Whitley, E., Jr. (1966), *Proc. Roy. Soc., Ser. B 166*, 124.
Shinoda, T., Titani, K., and Putnam, F. W. (1970), *J. Biol. Chem. 245*, 4463.
Terry, W. D., Page, D. L., Kimura, S., Isobe, T., Osserman, E. F., and Glenner, G. G. (1973), *J. Clin. Invest. 52*, 1276.
Titani, K., Shinoda, T., and Putnam, F. W. (1969), *J. Biol. Chem. 244*, 3550.
Titani, K., Wikler, M., Shinoda, T., and Putnam, F. W. (1970), *J. Biol. Chem. 245*, 2171.
Watanabe, S., and Hilschmann, N. (1970), *Hoppe Seyler's Z. Physiol. Chem. 351*, 1291.
Whitley, E. J., Jr. (1969), M.S. Thesis, Indiana University.
Whitley, E. J., Jr., Titani, K., and Putnam, F. W. (1969), *J. Biol. Chem. 244*, 3537.
Woods, K. R., and Wang, R. T. (1967), *Biochim. Biophys. Acta 133*, 369.

A Circular Dichroism Study of the Secondary Structure of Bradykinin†‡

John R. Cann,* John M. Stewart, and Gary R. Matsueda§

ABSTRACT: A systematic study has been made of the circular dichroism (CD) behavior of bradykinin, several of its analogs and peptide fragments, and model compounds. It is concluded that the secondary structure of bradykinin and its analogs is a time average of two interconverting structures—one disordered and one partially ordered due to a $3 \rightarrow 1$ type hydrogen bond bridging the Pro^7 residue. With increasing temperature the peptide spends a progressively greater fraction of its time intramolecularly hydrogen bonded until at 80° it spends all of its time in the partially ordered configuration. This transition is characterized by large positive values of $\Delta H°$ and $\Delta S°$, which is interpreted to mean that the structure of water plays a dominant role in determining the configuration of the peptide. Indeed, bradykinin possesses considerable conformational freedom and can assume different configurations in response to changes in solvent composition at constant temperature, *e.g.*, changing from water or buffer to 90% dioxane–water mixture. The nonadditivity of the CD spectra of peptide fragments and the comparative CD of bradykinin and its analogs in 90% dioxane are indicative of interactions between the amino acid residues in the bradykinin molecule.

The circular dichroism (CD) spectrum of bradykinin (the biologically active nonapeptide, Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg) in the wavelength interval 260–210 mμ, shows two weak bands, a negative one centered at 234 mμ and a positive one at 221 mμ. These spectral features suggested to Brady *et al.* (1971) that bradykinin has essentially a random-coil configuration in solution. More recently, however, the spectrum has been reinterpreted by one of us (Cann, 1972) as being indicative of a secondary structure having some order, most

† From the Department of Biophysics and Genetics and the Department of Biochemistry, University of Colorado Medical Center, Denver, Colorado 80220. *Received May 7, 1973.* Supported in part by Research Grants 5R01 HL 13909-21 and HE 12325 from the National Institutes of Health, U. S. Public Health Service, and a grant from the Population Council, New York. This publication is No. 534 from the Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado.

‡ The authors acknowledge the technical assistance of Robert O. Coombs, Kathie Hill, Martha Knight, and Roberta Tudor.

§ National Institutes of Health, Postdoctoral Fellow, 5 F02 AM-49277-02.