Horwitz, J., Strickland, E., and Billups, C. (1969), *J. Amer. Chem. Soc. 91*, 184.
Kabat, E. A. (1968), *Proc. Nat. Acad. Sci. U. S. 59*, 613.
Kartha, G., Bello, J., and Harker, D. (1967), *Nature 213*, 862.
Metzger, H., and Singer, S. (1963), *Science 142*, 674.
Millar, D. B., Frahali, V., and Willick, G. E. (1969), *Biochemistry 8*, 2416.
Miller, W. G., and Goebel, C. V. (1968), *Biochemistry 7*, 3925.
Porter, R. R. (1959), *Biochem. J. 73*, 119.
Press, E. M., and Hogg, N. M. (1969), *Nature (London) 223*, 807.
Pysh, E. S. (1966), *Proc. Nat. Acad. Sci. U. S. 56*, 826.
Ross, D. L., and Jirgensons, B. (1968), *J. Biol. Chem. 243*, 2829.
Sarkar, P. K., and Doty, P. (1966), *Proc. Nat. Acad. Sci. U. S. 55*, 981.
Steiner, L., and Lowey, S. (1966), *J. Biol. Chem. 241*, 231.
Steiner, R., Tippolot, R., Edelhoch, H., and Frahali, V. (1964), *Biopolym. Symp. 1*, 355.
Strickland, E. H., Horwitz, J., and Billups, C. (1969), *Biochemistry 8*, 3205.
Stryer, L., and Haugland, R. P. (1967), *Proc. Nat. Acad. Sci. U. S. 58*, 719.
Tanford, C. (1968), *Accounts Chem. Res. 1*, 161.
Townend, R., Kumosinski, T. F., Timasheff, S. N. (1967), *J. Biol. Chem. 242*, 4538.
Townend, R., Kumosinski, T. F., Timasheff, S. N., Fasman, G. D., and Davidson, B. (1966), *Biochem. Biophys. Res. Commun. 23*, 163.
Velick, S. F., Parker, C. W., and Eisen, H. N. (1960), *Proc. Nat. Acad. Sci. U. S. 46*, 470.
Weinryb, I., and Steiner, R. F. (1970), *Biochemistry 9*, 135.
Wikler, M., Kohler, H., Shinoda, T., and Putnam, F. W. (1969), *Science 163*, 75.
Wilkinson, J. M. (1969), *Biochem. J. 112*, 173.

The Covalent Structure of a Human γG-Immunoglobulin. VI. Amino Acid Sequence of the Light Chain*

Paul D. Gottlieb, Bruce A. Cunningham, Urs Rutishauser, and Gerald M. Edelman

ABSTRACT: The amino acid sequence of the constant region of the κ chain of the human γG-immunoglobulin, Eu, has been determined. This and previous studies establish the complete sequence of a light chain isolated from an intact serum immunoglobulin. With the exception of two residues, the sequence of the constant region is the same as that of constant regions of type K urinary Bence-Jones proteins which are Inv (3+). At position 108 the Eu light chain contains glycine instead of arginine. In addition, the Eu light chain contains aspartic acid at position 122 as has been found in two Bence-Jones proteins; in a third Bence-Jones protein, asparagine has been found at this position. The variable region of the Eu κ chain resembles variable regions of Bence-Jones proteins belonging to subgroup I. In view of the fact that it shows a larger amount of variation than other light chains of this subgroup, it may belong to a new subgroup.

Much of our present knowledge of antibody structure has been obtained through the study of immunoglobulin light chains which are readily obtained in the form of urinary Bence-Jones proteins (Edelman and Gally, 1962; for a recent review, see Edelman and Gall, 1969). Light chains contain a region of variable sequence and a region of relatively constant sequence (Hilschmann and Craig, 1965). The variable region has been implicated in the antigen binding function of antibody molecules (Singer and Thorpe, 1968). The constant (C_L) region of the light chain is presumably required to allow it to associate with the heavy chain so that V regions of both chains can form an antigen-combining site.

In the previous paper in this series (Cunningham *et al.*, 1968) we have presented the amino acid sequence of the variable region of the light chain of the human γGl myeloma protein Eu. We now report the complete amino acid sequence of the Eu light chain.

Materials and Methods

The techniques employed for isolation of light chains and CNBr fragments, and for gel filtration, ion-exchange chromatography, and amino acid analysis have been described previously (Edelman *et al.*, 1968; Waxdal *et al.*, 1968a,b; Cunningham *et al.*, 1968). Conditions for digestion with carboxypeptidase A, carboxypeptidase B, and leucine aminopeptidase have also been described (Gottlieb *et al.*, 1968). High-voltage paper electrophoresis was performed at pH 4.7 and 2.0 (Schwartz and Edelman, 1963).

Sequence Determinations Using the Dansyl-Edman[1] *Method.*

* From The Rockefeller University, New York, New York 10021. *Received December 8, 1969.* This work was supported in part by grants from the National Science Foundation (GB 8371) and the U. S. Public Health Service (AM 04256 and AI 09273).

[1] Abbreviations used that are not listed in *Biochemistry 5*, 1445 (1966), are: DNS and dansyl, 1-dimethylaminonaphthalene-5-sulfonyl; Asx, aspartic acid or asparagine; Glx, glutamic acid or glutamine.

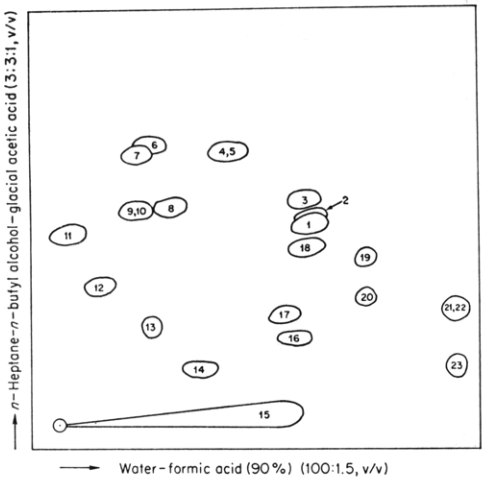

FIGURE 1: Sketch showing the relative positions of dansyl derivatives after two-dimensional chromatography on a polyamide layer. Chromatography was performed in the first dimension using water–formic acid (90%) (100:1.5, v/v). Chromatography in the second dimension was performed in *n*-heptane–*n*-butyl alcohol–glacial acetic acid (3:3:1, v/v). The symbol ⊙ in the bottom left corner denotes the point of sample application. Numbers designate dansyl derivatives. (1) DNS-NH_2; (2) *O*-DNS-tyrosine (yellow); (3) DNS-alanine; (4) DNS-valine; (5) DNS-proline; (6) DNS-isoleucine; (7) DNS-leucine; (8) DNS-methionine; (9) DNS-phenylalanine; (10) suspected to be α-DNS-ε-phenylthiocarbamyllysine, see text; (11) α-*O*-di-DNS-tyrosine (yellow); (12) α,ε-di-DNS-lysine; (13) suspected to be α-DNS-tyrosine, see text; (14) α-DNS-*S*-carboxymethylcysteine; (15) DNS-OH; (16) DNS-aspartic acid; (17) DNS-glutamic acid; (18) DNS-glycine; (19) DNS-threonine; (20) DNS-serine; (21) α-DNS-histidine; (22) ε-DNS-lysine; (23) DNS-arginine.

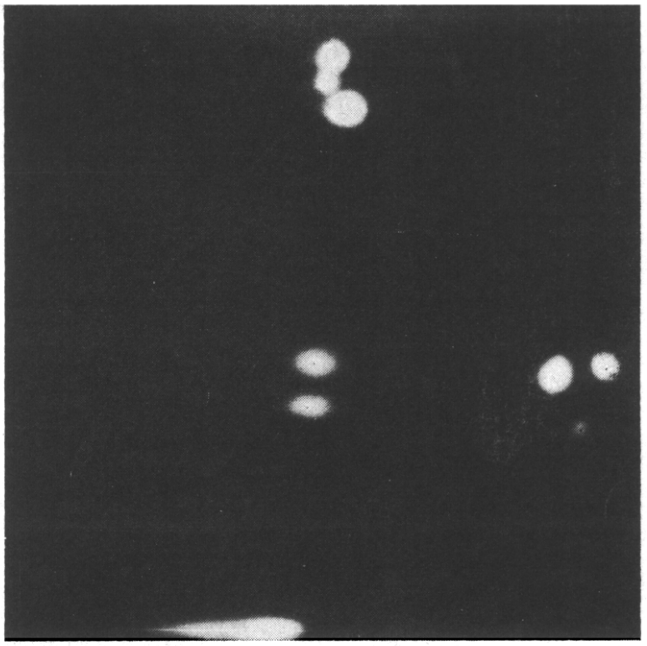

FIGURE 2: Photograph of a two-dimensional chromatogram used for the identification of dansylamino acids. The sample was applied at the lower left corner. The first dimension was run from left to right, and the three standards for the first dimension appear at the top of the chromatogram (from bottom to top, DNS-NH_2, DNS-glycine, DNS-alanine). The second dimension (*n*-heptane, *n*-butyl alcohol, and glacial acetic acid) was run from bottom to top, and the three standards for the second dimension appear at the extreme right of the chromatogram (from left to right, DNS-NH_2, DNS-glycine, and DNS-alanine). The two spots in the center of the chromatogram are DNS-NH_2 and DNS-glycine.

Sequence analysis was performed using the dansyl-Edman technique as described previously (Gray, 1967a,b; Cunningham *et al.*, 1968). Dansyl derivatives were submitted to thin-layer chromatography on polyamide sheets (Cheng Chin Trading Co., Taipei, Taiwan) using solvent 1 (200 ml of water and 3 ml of 90% formic acid) and solvent 3 (60 ml of *n*-heptane, 60 ml of *n*-butyl alcohol, and 20 ml of glacial acetic acid) of Woods and Wang (1967) for the first and second dimensions, respectively.

A sketch showing the relative positions of the various dansyl derivatives is presented in Figure 1. Dansylamide (spot 1) appears in all samples and serves as a convenient internal standard. In most instances the identification of dansylamino acids is made without difficulty by running a single two-dimensional chromatogram similar to that shown in Figure 2. However, several aspects of the determination of amino acid sequence by this method deserve further comment. *O*-DNS-tyrosine (Figure 1, spot 2) has a yellow fluorescence and frequently streaks in the second dimension, but this does not interfere with detection of DNS-aspartic acid, DNS-glutamic acid, and DNS-glycine (Figure 1, spots 16, 17, and 18, respectively). Occasionally DNS-alanine (Figure 1, spot 3) is not resolved from DNS-NH_2, and the chromatogram must be repeated. DNS-proline and DNS-valine must be distinguished by their different stabilities during acid hydrolysis, and the presence of a leucine or isoleucine must be confirmed by another method (Cunningham *et al.*, 1968). α-DNS-histidine is infrequently encountered, but when it is suspected in a peptide containing lysine, some care is required. Even after several steps of the Edman degradation, some material chromatographing in the position of ε-DNS-lysine is still detectable. Spot 13 (Figure 1) is assumed to be α-DNS-tyrosine because it lacks the yellow fluorescence characteristic of the *O*-DNS group, and it appears only when α-*O*-di-DNS-tyrosine (spot 11) is present. In the absence of phenylalanine, spot 10 (Figure 1) is assumed to be α-DNS-ε-phenylthiocarbamyllysine as it is often seen when a lysyl residue is encountered after several steps of the Edman degradation (Cunningham *et al.*, 1968). Hence it is often observed when the dansyl-Edman technique is used to establish the overlap of tryptic peptides. The amount of sample which must be applied is larger than usual, suggesting instability of this derivative during acid hydrolysis.

The chromatograms used in determining the amino acid sequence of peptide LT15a (Table I, see Figure 4) are shown in Figure 3 to exemplify the method. In addition to the carboxy-terminal lysine, this decapeptide contains only serine, threonine, and leucine, and hence is an example of a repeating sequence of amino acids. The amino acid sequence of peptide LT15a is easily deduced from the chromatograms (Figure 3). Several aspects of these chromatograms deserve comment. In addition to ε-DNS-lysine, DNS-serine, and DNS-NH_2, plate 1 (Figure 3) contains an additional fluorescent spot which does not correspond to the position of a dansylamino acid. The amino-terminal residue of a peptide is frequently contaminated with some reagent which will react with dansyl

chloride. At later steps, side products, presumably of the Edman degradation, often result in fluorescent spots which chromatograph beyond leucine in the second dimension (Figure 1). Peptides containing valine and isoleucine often yield dansyl-dipeptides, and these chromatograph in positions different from free dansylamino acids. None of these extra spots will interfere with the determination of the amino acid sequence of a peptide of known composition.

Examination of the latter portion of the sequence of peptide LT15a (Figure 3) shows that each residue is contaminated with the previous one, indicating that the Edman degradation is slightly out of phase. In several of the plates in Figure 3 (*e.g.*, plate 5), the level of background amino acid contamination can be clearly observed. In addition to DNS-serine resulting from incomplete Edman degradation, plate 5 contains traces of DNS-aspartic acid, DNS-glutamic acid, DNS-valine, and DNS-leucine, although the composition of the peptide showed none of these amino acids. These cause no difficulty, and when less sample is applied to the chromatogram (*e.g.*, plate 8), no background contamination is visible. In plate 10 (Figure 3) an exceptionally large application of sample allows detection of the spot suspected to be α-DNS-ε-phenylthiocarbamyllysine in the presence of a high level of background. Due to the great sensitivity of the dansyl method, care must be taken in reagent purification, glassware washing, and general manipulations to maintain a minimal level of background amino acid contamination.

It is clear that the dansyl-Edman method described above is qualitative and requires considerable care. In all cases where sequences have been checked using quantitative methods, the dansyl-Edman method has proved reliable, and it has been the sequence method of choice in this and the following papers of this series.

Tryptic Digestion of CNBr Fragment L_3. Fragment L_3 (217 mg) was suspended in 20 ml of distilled water at 37°. The pH was adjusted to 8.0, at which point the turbidity of the suspension increased. Trypsin (Calbiochem, Los Angeles, Calif., B Grade, Lot 65345) was added (3 mg, 1.5% w/w), and the pH was maintained at 8.0 using a Radiometer TTT1b titrator. The turbidity decreased considerably after 30 min, and after 6 hr the digest was lyophilized.

Chymotryptic Digestion of L_3. Fragment L_3 (91 mg) was suspended in 20 ml of distilled water at 37°, and the pH was adjusted to 8.0 with 1 N NH_4OH. Chymotrypsin (α-chymotrypsin, Worthington Biochemical Corp., Freehold, N. J., three-times recrystallized, CD1-6150-1) was added (0.9 mg, 1% w/w) to the cloudy suspension, and the pH was maintained at 8.0. After 30 min the solution was clear, and after 4 hr the digest was lyophilized.

Results

Previous communications in this series have described the isolation and ordering of the CNBr fragments of the Eu light chain (Waxdal *et al.*, 1968a), and the isolation of all of the tryptic peptides (Cunningham *et al.*, 1968). CNBr fragments L_1 and L_2 comprise residues 1–97 of the light chain. Fragments L_1 and L_2 and the first 11 residues of fragment L_3 constitute the variable (V_L) region of the Eu light chain, the sequence of which has been reported (Cunningham *et al.*, 1968). The remaining 106 residues of fragment L_3 constitute the constant (C_L) region, the amino acid sequence of which is discussed below.

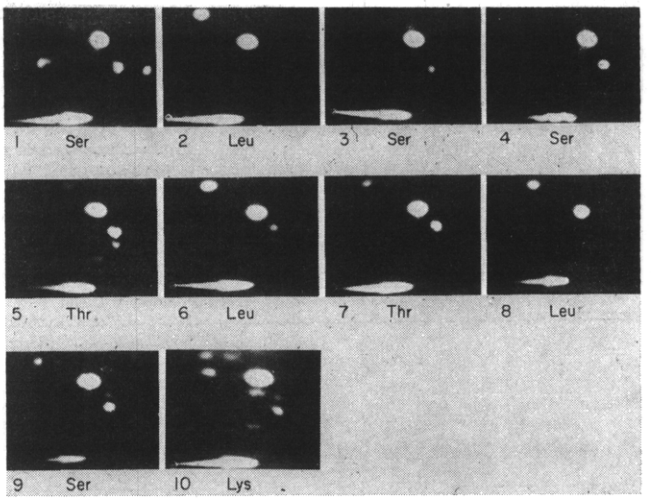

FIGURE 3: Chromatograms used in determining the amino acid sequence of the decapeptide, LT15a (Table I and Figure 4). The plates are numbered 1–10, and the amino acid identified is written directly below each plate. The point of sample application is at the lower left corner of each plate. The direction of chromatography is from left to right in the first dimension (water–formic acid) and from bottom to top in the second dimension (*n*-heptane–*n*-butyl alcohol–acetic acid). The standards run on each chromatogram (see Figure 2) were not photographed.

Tryptic Peptides. The isolation of peptides LT8–LT20 (Figure 4) from a tryptic digest of light chains has been described (Cunningham *et al.*, 1968). These peptides account for all of the residues of fragment L_3 and sequences obtained for these peptides are summarized in Figure 4.

In order to obtain additional material for sequence determination, peptides were also isolated from a tryptic digest of fragment L_3. Gel filtration of this digest on Sephadex G-25 (fine) in *n*-propyl alcohol–acetic acid–water (1:1:3, v/v) yielded three major fractions. Material from the first fraction was filtered on Sephadex G-50 to give one major and two minor fractions. The major fraction was submitted to ion-exchange chromatography on AGlX4 (Figure 5). Material from each of the fractions A, B, and D (Figure 5) was further purified by ion-exchange chromatography on AG50X4. Peptides LT15a, LT18a (Figure 4, Table I) and the NH_2-terminal tryptic peptide of fragment L_3 (residues 98–103) were obtained from fraction A. Fraction B yielded peptide LT18, and fraction D yielded peptides LT10 and LT10a. Fractions E and F both contained peptide LT14, and fraction H yielded peptide LT11.

The amino acid compositions of peptides LT10a, LT15a, and LT18a are presented in Table I, and their partial amino acid sequences are given in Figure 4. These peptides are the result of chymotryptic activity during the tryptic digest. Although the amino acid composition of peptide LT18a (Table I) indicates that this peptide is contaminated, the amino acid sequence as determined by the dansyl-Edman method is unambiguous. The presence of phenylalanine, aspartic acid, arginine, and of exactly 2.0 residues of serine (not low due to losses during hydrolysis) is probably due to contamination of peptide LT18a with a peptide spanning

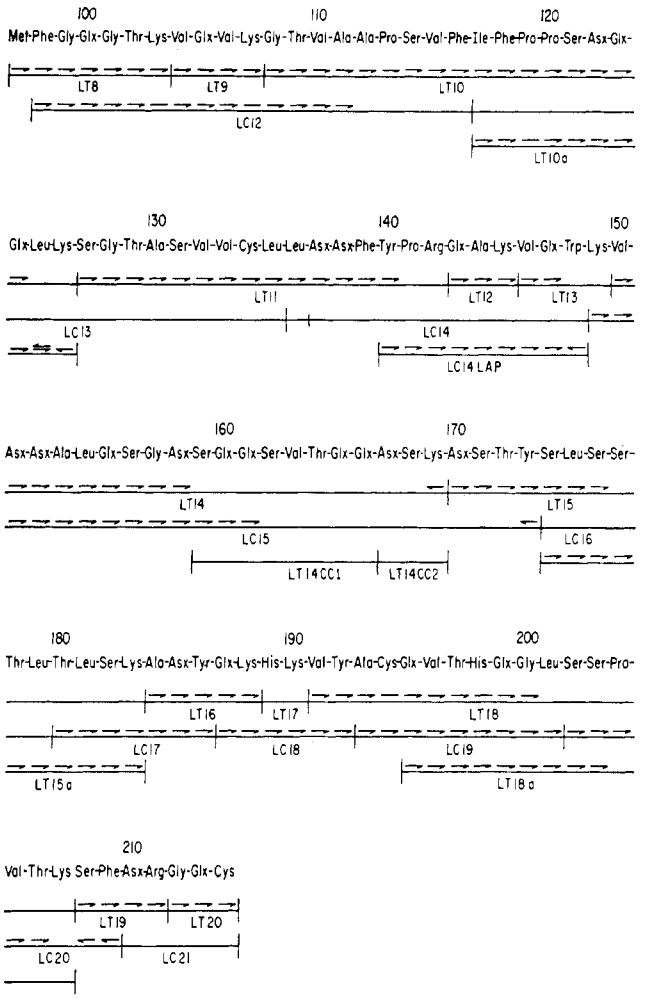

FIGURE 4: Amino acid sequence of fragment L_3. Residues are numbered from the NH_2 terminus of the light chain; fragment L_3 begins at Phe_{98}. Tryptic (LT), chymotryptic (LC), and other peptides used to establish the sequence are represented by solid lines. Peptides are numbered from the NH_2 terminus of the light chain. — represents the results of direct sequence analysis by the dansyl-Edman method. ⌐ represents the results of carboxypeptidase A or B digestions. —indicates that no dansylamino acid was detected. Refer to text for other nomenclature.

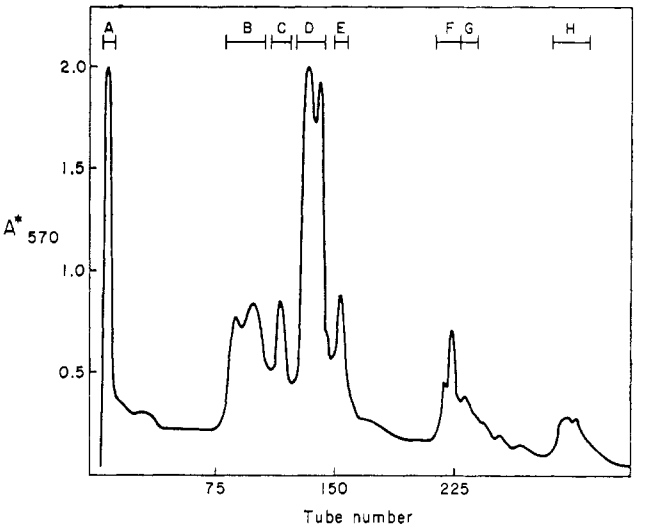

FIGURE 5: Fractionation of the large tryptic peptides of L_3 by ion-exchange chromatography on a column (0.9 × 60 cm) of AG1X4 at 38°. Peptides were eluted by the method of Schroeder (1967). Ninhydrin color yield on an arbitrary scale (A_{570}*) was determined automatically (Cunningham *et al.*, 1968). Each tube contained 2.9 ml of effluent.

LT18a and the succeeding tryptic peptide, LT19 (Figure 4), resulting from incomplete cleavage by trypsin at the lysine at position 207.

Chymotryptic Peptides. The chymotryptic digest of fully reduced and alkylated L_3 was filtered on a column of Sephadex G-25 (Figure 6). Peptide LC15 (Figure 4, Table I) was obtained from fraction A by successive gel filtrations on Sephadex G-50 and Sephadex G-25. Fraction B (Figure 6) was filtered on Sephadex G-25 and then fractionated by ion-exchange chromatography on AG50X4 to yield peptides LC12 and LC13. Fraction C (Figure 6) was also chromatographed on AG50X4 (Figure 7). Gel filtration of fraction B (Figure 7) on Sephadex G-25 yielded peptide LC16. Similar treatment of fraction H gave peptide LC20, and of fraction K gave peptides LC17 and LC21. Fractions F and Q (Figure 7) yielded peptides LC19 and LC14, respectively, without further purification. Peptide LC18 was not obtained from this digest, and was isolated from another chymotryptic digest of L_3 by high-voltage paper electrophoresis at pH 4.7.

The amino acid compositions of the chymotryptic peptides of fragment L_3 are presented in Table I.

Sequence of the Chymotryptic Peptides. The amino acid sequences obtained by dansyl-Edman analysis of chymotryptic peptides of L_3 are shown in Figure 4. Peptide LC14 (Figure 4, Table I) required special attention because it was a mixture of two nearly identical peptides, one of which contained an additional leucine at its NH_2 terminus. This mixture was treated with leucine aminopeptidase until amino acid analysis indicated no further release of amino acids. The digest was fractionated on Sephadex G-25, and material in the major fraction was further purified by ion-exchange chromatography on AG50X4, yielding peptide LC14-LAP (Table I). The amino acid sequence of LC14-LAP was determined by the dansyl-Edman method and carboxypeptidase A digestion (Figure 4).

Amino Acid Sequence of L_3. The amino acid sequence of the entire CNBr fragment L_3 is given in Figure 4. The order of peptides LT8, LT9, and LT10 was established by direct sequence analysis of L_3 using the dansyl-Edman method (Cunningham *et al.*, 1968), and was confirmed by the isolation of peptide LC12 (Figure 4 and Table I). The amino acid composition of peptide LC13 (Table I) established the order of peptides LT10 and LT11 (Figure 4). The amino acid sequence of peptide LC14-LAP established the order: LT11-LT12-LT13. The sequence of the first nine residues of peptide LT14 was determined by dansyl-Edman analysis (Figure 4). The sequence of this peptide was completed by analysis of peptides LT14CCl and LT14CC2 (Bennett *et al.*, 1970) which were isolated from a chymotrypsin C digest of LT14. Peptide LC15 had an NH_2-terminal lysine followed by a sequence identical with that of peptide LT14, establishing the order LT13-LT14. Treatment of peptide LC15 with carboxypeptidase A for 50 min released 0.4 residue of tyrosine;

TABLE I: Amino Acid Compositions of Peptides from L_3.[a]

	LC12	LC13	LC14	LC14-LAP	LC15	LC16	LC17	LC18	LC19	LC20	LC21	LT10a	LT15a	LT18a
Lys	2.2 (2)	1.0 (1)	1.1 (1)	1.1 (1)	1.9 (2)		1.1 (1)	2.1 (2)		1.0 (1)		1.0 (1)	1.0 (1)	1.0 (1)
His								1.1 (1)	1.0 (1)					0.7 (1)
Arg			1.1 (1)	1.0 (1)							1.0 (1)			0.3
CMCys		0.5 (1)							0.9 (1)		0.7 (1)			
Asp		1.0 (1)	1.8 (2)		4.6 (5)		1.0 (1)				1.0 (1)	1.1 (1)		0.3
Thr	1.9 (2)	1.0 (1)			1.8 (2)	1.0 (1)	0.9 (1)		0.9 (1)	0.9 (1)			1.9 (2)	1.9 (2)
Ser	1.0 (1)	2.3 (3)			3.6 (5)	2.8 (3)	1.1 (1)			2.4 (3)		1.1 (1)	3.4 (4)	2.0 (2)
Glu	2.2 (2)	2.0 (2)	2.0 (2)	2.0 (2)	5.1 (5)			1.0 (1)	2.0 (2)		1.0 (1)	2.0 (2)		1.9 (2)
Pro	1.3 (1)	1.8 (2)	0.9 (1)	0.9 (1)						1.0 (1)		2.0 (2)		1.1 (1)
Gly	2.7 (3)	1.1 (1)			1.2 (1)				1.0 (1)		1.0 (1)			1.4 (1)
Ala	1.9 (2)	0.9 (1)	0.9 (1)	1.1 (1)	1.1 (1)		1.1 (1)		1.0 (1)					0.4
Val	3.9 (4)	1.8 (2)	1.0 (1)	1.1 (1)	1.8 (2)			1.0 (1)	1.1 (1)	1.0 (1)				1.9 (2)
Ile		0.8 (1)										0.8 (1)		
Leu		1.5 (2)	0.4		1.1 (1)	2.2 (2)	0.9 (1)		1.0 (1)			1.0 (1)	3.0 (3)	1.1 (1)
Tyr			1.0 (1)	0.9 (1)	0.8 (1)		0.7 (1)	0.8 (1)						
Phe	1.9 (2)	1.1 (1)	1.0 (1)							0.9 (1)		0.9 (1)		0.2
Trp			(1)	(1)										
Total residues	19	19	12	9	25	6	7	6	9	8	5	10	10	13
Yield (%)[b]	38	28	33	17[c]	20	48	14	31	57	45	23	16	15	3

[a] Values reported are amino acid residues. Amino acids present at a level of 0.1 residue or less are omitted. Values in parentheses represent assumed integral numbers of residues. [b] Yields are based on micromoles of peptides isolated compared with the micromoles of protein originally digested. [c] Yield is based on micromoles of LC14 digested with leucine aminopeptidase.

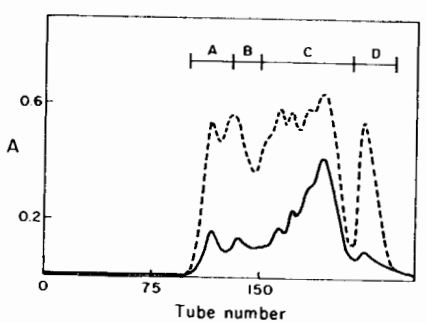

FIGURE 6: Initial fractionation of the chymotryptic peptides of L_3 by gel filtration on a column (2.5 × 157 cm) of Sephadex G-25 (fine) in *n*-propyl alcohol–acetic acid–water (1:2:97, v/v). The solid line indicates absorbance at 280 mμ, and the dashed line represents absorbance at 230 mμ. Each tube contained 2.8 ml of effluent.

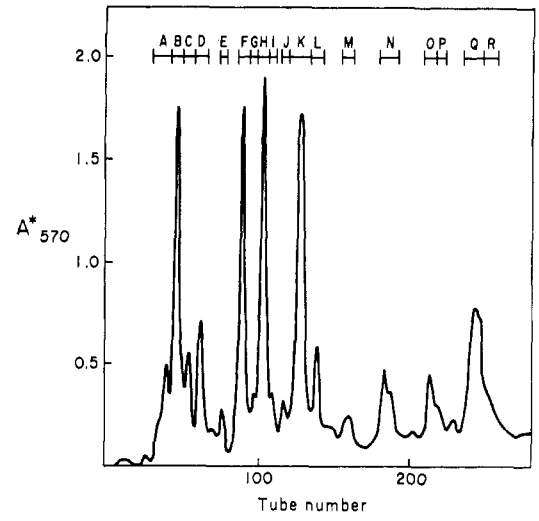

FIGURE 7: Ion-exchange chromatography of fraction C (Figure 6) on a column (0.9 × 13 cm) of AG50X4 at 37°. Peptides were eluted by means of a linear gradient of 50 ml each of 0.05 M pyridinium-acetate (pH 3.1) and 0.2 M pyridinium-acetate (pH 3.1). This was followed by a second linear gradient of 150 ml each of 0.2 M pyridinium-acetate (pH 3.1) and 1.0 M pyridinium-acetate (pH 5.6). The column was subsequently washed with 2.0 M pyridinium-acetate (pH 5.6). The solid line indicates ninhydrin color yield on an arbitrary scale (A_{570}*) determined automatically (Cunningham *et al.*, 1968). Each tube contained 1.5 ml of effluent.

this finding and the amino acid composition of LC15 (Table I) established the order LT14-LT15 (Figure 4). Determination of the amino acid sequences of peptides LC17 and LC18 by the dansyl-Edman method established the order of peptides LT15–LT18. Similar analysis of peptide LC20 and the amino acid composition of LC21 completed the ordering of the tryptic peptides of L_3 (Figure 4).

The chymotryptic peptides LC12–LC21 account for all the residues of CNBr fragment L_3, and their sequences and compositions provide unequivocal evidence for the order of the tryptic peptides LT8–LT20.

Discussion

The complete amino acid sequence of the Eu light chain is presented in Figure 8. The determination of the positions of asparaginyl and glutaminyl residues is reported in another paper in this series (Bennett *et al.*, 1970).

We have already discussed the relationship of the V_L region of the Eu light chain to V regions of Bence–Jones proteins (Cunningham *et al.*, 1968). Several workers (Hood *et al.*, 1967; Milstein, 1967; Niall and Edman, 1967) have suggested that V_κ regions can be divided into three subgroups which differ in length, and Milstein (1969) has proposed a basic sequence which underlies each subgroup. The Eu V region is identical in length with subgroup I. Moreover, the sequence of the V_κ region of Eu is consistent with the basic sequence of subgroup I at 85% of the positions, as contrasted with 64% and 70% for subgroups II and III, respectively. It has been suggested (Milstein, 1969; Hood and Talmage, 1970) that subgroup I can be further divided to create a fourth subgroup which would include the Eu light chain.

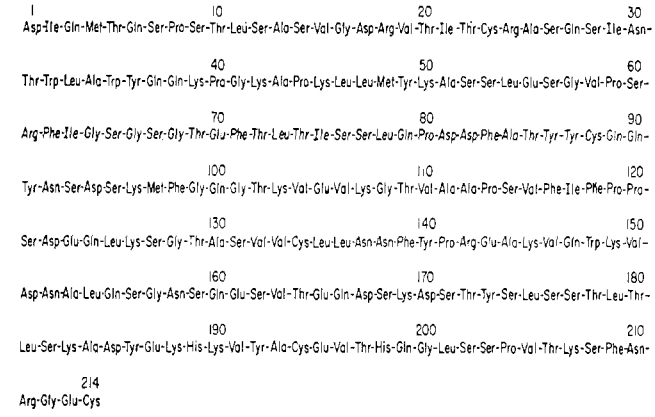

FIGURE 8: The complete amino acid sequence of the Eu light chain.

Residues 39–42 of the Eu light chain, of protein Roy (Hilschmann, 1967), and of protein Ag (Putnam, 1969; Titani *et al.*, 1969), all of which are considered to belong to subgroup I, have the same amino acid composition but differ in sequence. The sequence of Eu light chain in this region is Lys-Pro-Gly-Lys, and that of proteins Roy and Ag is Gly-Pro-Lys-Lys. The sequence Lys-Pro-Gly-Lys in this region of the Eu light chain has been confirmed by leucine aminopeptidase digestion of the peptide Tyr-Gln-Gln-(Pro,-Gly,Lys_2) followed by gel filtration of the digest on Sephadex G-25 and sequence analysis of the resulting tetrapeptide (Pro,Gly,Lys_2) by the dansyl-Edman procedure.

There are a number of variations in the C region of the Eu light chain which deserve mention. Residue 108 of this chain is glycine, but in every other κ chain yet studied it is arginine. The glycyl residue may represent a rare allotypic variation in the C region and thus cannot as yet be assigned with certainty to the V region. The Eu light chain contains aspartic acid at position 122 in agreement with Bence–Jones protein Roy (Hilschmann, 1967) and a peptide isolated from Bence–Jones protein Rad (Milstein, 1966). This differs from Bence–Jones protein Ag which has been reported to contain asparagine at this position (Titani *et al.*, 1969). In accord with other studies (Baglioni *et al.*, 1966), protein Eu is Inv(3+) and contains valine at position 191.

If it is confirmed that the Eu κ chain belongs to subgroup I, then it shows more variation from the basic sequence than any κ Bence–Jones protein (Cunningham *et al.*, 1968). If other κ chains are found with V regions which resemble the Eu light chain much more closely than subgroup I, then this may well provide a basis for designating a fourth subgroup of κ chains. In any case, the present studies of a light chain from a myeloma protein show that there is no *fundamental* structural difference between Bence–Jones proteins and light chains of serum immunoglobulins.

Acknowledgments

We are grateful to Mrs. Helvi Hjelt and Miss Joan Low for their valuable technical assistance.

References

Baglioni, C., Zonta, L. A., Cioli, D., and Carbonara, A. (1966), *Science 152*, 1517.

Bennett, C., Konigsberg, W. H., and Edelman, G. M. (1970), *Biochemistry 9*, 3181.
Cunningham, B. A., Gottlieb, P. D., Konigsberg, W. H., and Edelman, G. M. (1968), *Biochemistry 7*, 1983.
Edelman, G. M., and Gall, W. E. (1969), *Annu. Rev. Biochem. 38*, 415.
Edelman, G. M., Gall, W. E., Waxdal, M. J., and Konigsberg, W. H. (1968), *Biochemistry 7*, 1950.
Edelman, G. M., and Gally, J. A. (1962), *J. Exp. Med. 116*, 207.
Gottlieb, P. D., Cunningham, B. A., Waxdal, M. J., Konigsberg, W. H., and Edelman, G. M. (1968), *Proc. Nat. Acad. Sci. U. S. 61*, 168.
Gray, W. R. (1967a), *Methods Enzymol. 11*, 139.
Gray, W. R. (1967b), *Methods Enzymol. 11*, 469.
Hilschmann, N. (1967), *Hoppe-Seyler's Z. Physiol. Chem. 348*, 1077.
Hilschmann, N., and Craig, L. C. (1965), *Proc. Nat. Acad. Sci. U. S. 53*, 1403.
Hood, L., Gray, W. R., Sanders, B. G., and Dreyer, W. J. (1967), *Cold Spring Harbor Symp. Quant. Biol. 32*, 133.
Hood, L., and Talmage, D. W. (1970), *Science 168*, 325.
Milstein, C. (1966), *Biochem. J. 101*, 352.
Milstein, C. (1967), *Nature (London) 216*, 330.
Milstein, C. (1969), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 2*, 301.
Niall, H. D., and Edman, P. (1967), *Nature (London) 216*, 262.
Putnam, F. W. (1969), *Science 163*, 633.
Schroeder, W. A. (1967), *Methods Enzymol. 11*, 361.
Schwartz, J. H., and Edelman, G. M. (1963), *J. Exp. Med. 118*, 41.
Singer, S. J., and Thorpe, N. O. (1968), *Proc. Nat. Acad. Sci. U. S. 60*, 1371.
Titani, K., Whitley, E. J., and Putnam, F. W. (1968), *J. Biol. Chem. 244*, 3521.
Waxdal, M. J., Konigsberg, W. H., and Edelman, G. M. (1968a), *Biochemistry 7*, 1967.
Waxdal, M. J., Konigsberg, W. H., Henley, W. L., and Edelman, G. M. (1968b), *Biochemistry 7*, 1959.
Woods, K. R., and Wang, K.-T. (1967), *Biochim. Biophys. Acta 133*, 369.

The Covalent Structure of a Human γG-Immunoglobulin. VII. Amino Acid Sequence of Heavy-Chain Cyanogen Bromide Fragments H_1–H_4*

Bruce A. Cunningham, Urs Rutishauser, W. Einar Gall, Paul D. Gottlieb, Myron J. Waxdal, and Gerald M. Edelman

ABSTRACT: The amino acid sequence of the cyanogen bromide fragments H_1–H_4 from the γ chain of the immunoglobulin Eu has been determined. These four fragments contain the amino-terminal 252 residues of the heavy chain and include the entire Fd(t) portion of the molecule. Comparisons of the sequence with that of the κ chain from the same protein and with the reported sequences of portions of other immunoglobulins indicate that: (1) heavy chains, like light chains, are composed of a variable region and a constant region. (2) The variable region of γ chains (V_γ) begins at the amino terminus and extends for at least 114 residues; thus the V_L and V_H regions are similar in length. (3) There are at least two subgroups of heavy-chain variable regions. (4) The variable region of the Eu heavy chain (V_H) is homologous to the variable region of the light chain (V_L), but there is no obvious, special relationship between the V_H and V_L regions from the same immunoglobulin molecule. (5) The beginning of the constant (C_H) region (C_H1, residues 119–220) of the heavy chain is homologous in sequence to the constant region of the Eu light chain (C_L).

Studies on the amino acid sequences of light chains (Hilschmann and Craig, 1965; Titani *et al.*, 1967; Cunningham *et al.*, 1968; Gottlieb *et al.*, 1970) support the conclusion that they are composed of regions of variable and constant amino acid sequences. A variety of data (Frangione and Franklin, 1965; Press and Piggot, 1967; Gottlieb *et al.*, 1968) suggested that heavy chains also have variable (V) and constant (C) regions. Confirmation of this hypothesis requires the determination of the sequences of several immunoglobulin heavy chains.

In this and the subsequent paper in this series (Rutishauser *et al.*, 1970), we present the determination of the complete amino acid sequence of the γ chain (446 residues) from protein Eu. All of the residues of the heavy chain of this immunoglobulin are accounted for by CNBr fragments H_1–H_7 (Waxdal *et al.*, 1968b). Here we report the amino acid sequence of CNBr fragments H_1–H_4 which comprise the first 252 residues

* From The Rockefeller University, New York, New York 10021. *Received December 8, 1969.* This work was supported in part by grants from the National Science Foundation (GB 8371) and the U. S. Public Health Service (AM 04256 and AI 09273). Preliminary reports of these results have appeared elsewhere (Gottlieb *et al.*, 1968; Edelman *et al.*, 1969).