The Amino-Terminal Sequence of the Heavy Chain of Human Immunoglobulin M*

J. Claude Bennett†

ABSTRACT: In the study of the heavy chain from four IgM immunoglobulins no α-amino-terminal group was detected by fluorodinitrobenzene, cyanate, or dimethylaminonaphthalenesulfonyl chloride. The N-terminal residue appears to be a pyrrolidonecarboxylic acid in all four cases.

The N-terminal peptide from each of the four has been isolated and a partial sequence determination carried out. The sequences extending over five to nine residues reveal a structural homology among these μ chains and suggest homology with other immunoglobulin chains.

There is now considerable evidence which suggests that the specific affinity of an antibody for its antigen is due to the tertiary conformation of the binding site on the antibody molecule. The experiments of Haber (1964) and of Whitney and Tanford (1965) would indicate that this in turn depends upon the amino acid sequence of the constituent polypeptide chains. It now seems established that both the light (L) and the heavy (H) chains of immunoglobulin G (IgG)[1] contain an N-terminal sequence of a limited variability (Hilschman and Craig, 1965; Porter, 1967). In the case of the L chain this variable part extends over approximately half of the total length of the chain. The extent of variability within the H chain is not yet conclusively established, but it may be over a comparable distance. It is assumed, therefore, that this variable sequence at the amino-terminal end must be related in some way to the specific conformation making up the antibody binding site. Although data regarding the chain location of the binding are not uniform, it would seem that most of the activity resides in the H chain with perhaps some small part, either direct or indirect, being contributed by the L chain. At any rate, the concept has arisen that the N-terminal portion of the polypeptide chains contains a variable sequence and is responsible for forming the combining site (Singer and Doolittle, 1966).

It is assumed that the nature of the variability will be important in understanding the nature of the molecular configuration of the combining site. In addition, there seems to be good precedence from comparative amino acid sequences in other types of molecules (Bennett and Dreyer, 1964; Margoliash, 1963) that would indicate that a study of sequences in this part of the molecule from the standpoint of evolution and phylogeny may give us some indication as to the genetic mechanisms involved in producing the variability. With these goals in mind, we have begun to study the selected sequence at the amino-terminal portion of the heavy chain of the IgM molecule. This class of immunoglobulins is considered by many individuals to be an early response both from the standpoint of phylogeny (Marchalonis and Edelman, 1966) and perhaps ontogeny (Good and Papermaster, 1964).

Materials and Methods

Pathological IgM. These molecules were isolated from the plasma of four patients with a diagnosis of Waldenström's macroglobulinemia. Euglobulin preparations were obtained by precipitation of plasma with a 15-fold excess of cold water. The precipitate was resuspended in saline and reprecipitated three times. Any associated IgG which remained after the procedure was then removed by gel filtration on Sephadex G-200 in 1 M acetic acid. The IgM was then found to be homogeneous as evaluated by its pattern in the ultracentrifuge and in acrylamide gel disc electrophoresis (J. C. Bennett, unpublished data). These preparations gave a single line on agar diffusion against anti-IgM antisera. No IgG or IgA was detected when tested against appropriate antisera.

Enzymes. Chymotrypsin and pepsin were obtained from the Worthington Biochemical Corp. of Freehold, N. J. Subtilisin was obtained from the Sigma Chemical Corp., St. Louis, Mo., and Pronase from the K & K Laboratories, New York, N. Y. Chymotrypsin and subtilisin were stored in 0.001 N HCl at a concentration of 10 mg/ml. Pepsin was made up fresh in 4% formic acid immediately before each experiment and Pronase was dissolved in 0.001 N HCl before each experiment. Digestions with chymotrypsin, subtilisin, and Pronase were done at pH 8.2 in a pH-Stat. Digestion with pepsin was carried out in 4% formic acid.

Reagents and Other Materials. Guanidine hydrochloride was obtained in a highly purified form from

* From the Division of Rheumatology, Department of Medicine, University of Alabama in Birmingham, The Medical Center, Birmingham, Alabama. *Received June 10, 1968.* Supported in part by grants from the National Institutes of Health (GM 12456) and the National Science Foundation.

† The author is recipient of Research Career Development Award 5 K3 25386 and is a Markle scholar in academic medicine.

[1] Nomenclature of immunoglobulins is according to *Bull. World Health Organ. 30,* 447 (1964). Dansyl refers to dimethylaminonaphthalenesulfonyl chloride derivatives.

TABLE I: N-Terminal Sequences of IgM H Chains.[a]

Source	Digest	Sample Wt (mg) and Recov (%)	Sequence
Dos	Pepsin	40 and 20	Glp-Ser-Val-Ala-Asx-(Gly-Thr-Leu)Glx
	Chy	50 and 60	Glp-Ser-Val-Ala-Asx-(Gly-Thr-Leu)
	Sub	50 and 58	Glp-Ser-Val-Ala-Asx
	H^+		-Ser-Val-Ala-Asx
Bus	Chy	40 and 55	Glp-Ser-Val-Leu-Asx-Gly
	Sub	10 and 65	Glp-Ser-Val-Leu-Asx
Bal	Chy	50 and 60	Glp-Ser-Val-Ala-Glx-(Gly-Leu)
	Sub	50 and 65	Glp-Ser-Val-Ala-Glx
Dau	Chy	100 and 52	Glp-Ser-Val-Leu-Asx-(Ala-Thr)
	Sub	50 and 60	Glp-Ser-Val-Leu
	Pro	100 and 62	Glp-Ser

[a] Sequences of the N-terminal peptides from various enzyme digests of four IgM H chains are given. Chy, chymotrypsin; Sub, subtillisin; H^+, partial acid hydrolysis of the Dos subtilisin pepide; Pro, Pronase.

the Mann Research Laboratories, New York, N. Y. The optical density of 5 M guanidine hydrochloride solutions was less than 0.02 at 280 mμ. Sephadex for gel filtration was obtained from the Pharmacia Co. Dowex 50 for peptide purification was obtained from the Bio-Rad Corp., Los Angeles, Calif. Dimethylaminonaphthalenesulfonyl chloride was obtained from the California Corp. for Biochemical Research, Los Angeles, Calif. Iodoacetamide was obtained from Eastman Chemical Corp., Rochester, N. Y., and was recrystallized from hot H_2O. Phenyl isothiocyanate (Eastman) was redistilled *in vacuo* at 66–67°. All other materials used were reagent grade according to ACS standards. All solvents used in the Edman procedure were redistilled.

Purification of H Chains. The preparations of IgM were dissolved in 7 M guanidine hydrochloride at a concentration of 10 mg/ml, with the solution buffered at pH 7.6 with 0.5 M Tris-HCl. Reduction was allowed to proceed in the presence of 0.2 M β-mercaptoethanol at 37° for 4 hr and was followed by alkylation at room temperature for 15 min with a slight molar excess of iodoacetamide. The sample was subsequently dialyzed against water and freeze dried. Peptide chains were then separated on Sephadex G-200 in 5 M guanidine hydrochloride by the reverse-flow method of Small and Lamm (1966). The columns measured 100 × 2.5 cm and each sample load consisted of approximately 100 mg of protein.

Isolation of Acidic Peptides. Fully reduced and alkylated H chain was digested with one of the above enzymes at an enzyme:protein ratio of 1:50 (w/w) at 37° in the cell of a pH-Stat for 4 hr. A 50-mg digest was applied to a column (10 × 0.7 cm) of Dowex 50-X2 in the H^+ form. Four 5-ml water-wash fractions were col-

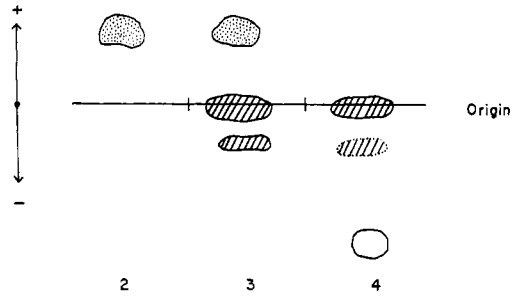

FIGURE 1: Drawing to show electrophoretic migration on paper of peptides obtained from water wash fractions from Dowex 50-X2. See text for methods. Fractions 2 and 3 contained the blocked N-terminal peptide which was ninhydrin negative (dotted). Fractions 3 and 4 contained ninhydrin-positive glycopeptides (cross-hatched). In addition fraction 4 contained a ninhydrin-positive spot (open circle) which gave a blank analysis following elution and was assumed to be a contaminant from the Dowex. Results here are at pH 3.5, 3500 V for 1 hr.

lected, pooled, and then recycled onto the same Dowex column and again 5-ml fractions were collected following water elution. The tubes were then frozen on the slant and dried. Elution fractions were spotted on paper (Whatman No. 3MM) and electrophoresed at pH 3.5, 4.4, or 6.5 at 3500 V for 1 hr; peptides were then eluted from paper following detection by ninhydrin (Katz *et al.*, 1959) or a starch–iodide (Pan and Dutcher, 1956) staining of a parallel strip of paper.

Electrophoresis of Peptides. Some procedures were carried out in the Gilson electrophorator under Varsol, while others were done using the Savant flat-plate system. Operating temperature in all cases was 22°.

Amino Acid Composition. Analyses were performed on the Beckman Model 120C automatic amino acid analyzer using 55-and 8-cm column separations for neutral and acidic and basic amino acids, respectively (Hubbard, 1965). Composition of small peptides was determined by the high-voltage paper electrophoretic technique of Dreyer and Bynum (1967).

Edman Degradation. For peptide sequencing the technique developed by Edman using phenyl isothiocyanate (Edman, 1956), as adapted by Gray (1967a,b) for utilization with the dansyl reagent, was followed. Although this gives only a qualitative test for the presence of each residue, in the cases utilizing up to six cycles the answers are quite unambiguous and good yields are obtained for each of the amino acids.

Identification of the dansyl derivatives is routinely done employing two techniques: (1) the electrophoretic system of Gray (1967a,b) utilizing the Savant flat plate at pH 4.5 and/or 1.9 as well as (2) the chromatographic system of Wilkinson as described by Press *et al.* (1966), in which the chromatograms were developed with chloroform–acetic acid (7:3) and then secondarily developed in the same dimension with methanol–aqueous ammonia (9:1). These two approaches afford a check on the qualitative determination for each dansyl derivative, and good agreement was obtained.

Partial Acid Hydrolysis. The terminal peptide obtained from a subtilisin digest of the Dos μ chain was subjected to partial acid hydrolysis. This was carried out in 2 N HCl at 100° for 10 min in a sealed tube (see results in Table I).

Results

Only trace amounts of amino acids with a free α-amino group could be detected in the heavy chain of human IgM by use of the fluorodinitrobenzene method (Sanger and Tuppy, 1951) or carbamylation (Stark, 1967). Dansylation on the whole protein carried out in 8 M urea failed to reveal even trace amounts of an N-terminal residue. Wilkinson *et al.* (1966) have demonstrated that the H chain of IgG contains a pyrrolidonecarboxylic acid residue at the N terminus. It was, therefore, felt that possibly the same structure might be present in the IgM H chain. Methodology designed to isolate peptides containing such a residue was undertaken employing various enzyme digests.

Peptides which contained no free amino groups were isolated by elution from a small column (10 × 0.7 cm) of Dowex 50-X2 (H^+ form) which had previously been equilibrated with water. The column was washed with four volumes of water and the pass-through volume collected in four fractions as described above. Peptides containing free amino groups were retained on the column. The third and fourth samples in each case contained two very faintly staining ninhydrin-positive peptides, which moved only slightly off the origin toward the cathode at pH 3.5. These peptides on elution and analysis were found to contain glucosamine and were therefore considered to be glycopeptides that had not been fully adsorbed on the column. Their yields were generally from 20 to 30% of calculated molar values and they were not further analyzed. Tubes 2 and 3 in each case contained a ninhydrin-negative peptide which did stain a deep blue with the starch–iodide method and moved toward the anode (see Figure 1). These were therefore considered to be the peptides which lacked the free α-amino group. At other pH values (4.4 and 6.5) these peptides demonstrated their expected acidic properties. In preparative experiments unstained peptides were eluted following identification by staining of parallel detection strips. This same methodological approach was used for all four of the IgM H chains and for each of the enzyme digests that was utilized. Although recovery of peptides on elution from paper is quite variable, the quantifications of the peptides calculated directly from compositional analyses are listed in Table I. Owing to the tendency for pyrrolidonecarboxylic acid to form as a result of cyclization of an internal glutamine residue, the possibility existed that the isolated peptide might actually be due to an artifact (Hirs *et al.*, 1956). The acidic conditions of the column would enhance this possibility. Hence a Pronase digestion was carried out on the pH-Stat and was followed immediately by the addition of fluorodinitrobenzene to block all free amino groups. The reaction mixture was acidified and most of the DNP peptides were extracted with ethyl acetate and ether. Additional DNP peptides were then removed by passage through a small column (1 × 3 cm) of talc in 1 M acetic acid. The pass-through volume was collected,

TABLE II: Comparative N-Terminal Sequences of Ig Chains.[a]

Chain	Source	References	Sequence
μ	Dos	This paper	Glp-Ser-Val-Ala-Asx-
	Bus	This paper	Glp-Ser-Val-Leu-Asx-
	Bal	This paper	Glp-Ser-Val-Ala-Glx-
	Dau	This paper	Glp-Ser-Val-Leu-Asx-
γ	Daw	Press *et al.* (1966)	Glp- Val-Thr-Leu-Arg-
	Cor	Press and Piggot (1967)	Glp- Val-Thr-
	Rabbit	Wilkinson *et al.* (1966)	Glp-Ser-Val-Glu-Glu- 50%
λ	Bo	Putnam (1967)	Glp-Ser-Ala-Leu-Thr-
	Ha	Putnam (1967)	Glp-Ser-Val-Leu-Thr-
	Sh	Putnam (1967)	Ser-Glu-Leu-Thr-
	HBJ 11	Hood *et al.* (1967)	Glp-Ser-Val-Leu-Thr-
	HBJ 8	Hood *et al.* (1967)	Glp-Ser-Ala-Leu-Ala-
	HBJ 15	Hood *et al.* (1967)	Glp-Ser-Ala-Leu-Thr
	HBJ 2	Hood *et al.* (1967)	Glp-Ser-Ala-Leu-Thr-
	HBJ 7	Hood *et al.* (1967)	Glp-Ser-Val-Leu-Thr-

[a] These sequences on human proteins are obtained from pathological samples. The rabbit sequence reflects the major peptide from pooled IgG H chains (γ chains).

freeze dried, resuspended in water, and applied to Dowex 50-X2. Following this procedure the same peptide was isolated, and it was concluded that these peptides must be derived from the N-terminal sequence.

The N-terminal peptides produced by enzymatic digestion of each of the immunoglobulin H chains according to each enzyme technique were purified from Dowex 50 and paper as described above. No difficulty was encountered with migration on paper even with the larger peptide derived from pepsin digestion. In each of the cases a variable amount of one or two glycopeptides was detected. A sequence study of these peptides has not yet been completed. The yields of the N-terminal peptides varied from 20 to 65% on an expected molar basis. Following isolation of the peptides from paper, amino acid compositions were determined on an aliquot. An additional aliquot, equivalent to the yield from 10 mg of H chain, was subjected to hydrolysis with dilute alkali in order to open the pyrrolidone ring (Dekker *et al.*, 1949). This was accomplished by exposure of the peptide to 1 N NaOH at room temperature for 16 hr. Following this hydrolytic procedure the sample tubes were placed in an environment of carbon dioxide (Dry Ice box) for neutralization. A subset aliquot was then ready for a first-step dansylation (Gray, 1967a,b). In each instance this revealed an unambiguous N-terminal glutamic acid residue. Remaining unblocked peptide was neutralized with dilute hydrochloric acid and desalted by passage through a polyacrylamide (P-2) column in water. This resulted in a salt-free solution of the peptide which was then dried and subjected to sequential Edman degradation (Gray, 1967a,b). Table I gives the results of the amino acid sequence determinations by the Edman procedure.

Discussion

This report presents evidence for a pyrrolidonecarboxylic residue at the N terminus of the heavy chain of pathological human IgM proteins. The presence of this blocked group has facilitated the isolation of the N-terminal peptides from these chains and has allowed their one-step purification for sequence determination.

The yields of the peptides on a molar basis after they have been through Dowex columns and eluted from paper would indicate a per cent recovery that is compatible with essentially a homogeneous N-terminal sequence for each of the pathological IgM H chains. The sequence data which extend over five to nine residues among the four chains would indicate that there is variation in the N terminus of the IgM H chains, but that the variation may be limited to only a few sites. This is the same situation as has been described in the variable part of the L chains (Putnam, 1967).

It has been found that the pyrrolidonecarboxylic acid group is present in the N terminus of the H chain of IgG (Wilkinson *et al.*, 1966). It is also present in human λ-type L chains (Hood, 1966) and has recently been described by Rüde and Givol (1968) for the L chains of rabbit. Comparative N-terminal sequences in the γ, μ, and λ chains are given in Table II. From these sequences a genetic relationship among the various immunoglobulin polypeptide chains is suggested. The first three residues (Glp-Ser-$^{\text{Ala}}_{\text{Val}}$) hold throughout with the exception of a supposed deletion of Ser in the γ chain and the absence of Glp in position 1 and substitution of Glu in position 3 for the Sh λ chain.

Although no conclusions can be made until much

longer sequences are established, present evidence on the H chain of IgM would indicate that there is variability in the N-terminal end of this chain, and it may be assumed that this relates in some way to antibody specificity. The relationship between specificity and sequence may be difficult to obtain due to the expected variation in sequences of any heterogeneous antibody preparation. However, the study of a homogeneous antibody such as has now been found among several pathological IgM proteins (Metzger, 1967) would be of interest. Two of the IgM proteins in this report (Bus and Bal) have antibodylike activity against human IgG (R. Schrohenloher, personal communication). Interpretation of the relationship of sequence data and structural specificity still presents an extraordinarily difficult problem.

Acknowledgments

The author is indebted to Thomas Hose, Robert Kelly, and Mary Hurst for excellent technical assistance and to Drs. Ralph Schrohenloher, Robert Stroud, and Leroy Hood for many helpful discussions. Plasma for these studies was obtained from patients of Dr. William Hammack in the Clinical Research Center, University Hospital (supported by U. S. Public Health Service Grant 2M01-FR-32-08). One sample, Dau, was supplied by Dr. John Harter, Peter Bent Brigham Hospital, Boston, Mass.

References

Bennett, J. C., and Dreyer, W. J. (1964), *Ann. Rev. Biochem. 33*, 205.

Dekker, C. A., Stone, D., and Fruton, J. S. (1949), *J. Biol. Chem. 181*, 719.

Dreyer, W. J., and Bynum, E. (1967), *Methods Enzymol. 11*, 32.

Edman, P. (1956), *Acta Chem. Scand. 10*, 761.

Good, R. A., and Papermaster, B. W. (1964), *Advan. Immunol. 4*, 1.

Gray, W. R. (1967a), *Methods Enzymol. 11*, 469.

Gray, W. R. (1967b), *Methods Enzymol.* 139.

Haber, E. (1964), *Proc. Natl. Acad. Sci. U. S. 52*, 1099.

Hilschman, N., and Craig, L. C. (1965), *Proc. Natl. Acad. Sci. U. S. 53*, 1403.

Hirs, C. H. W., Stein, W. H., and Moore, S. (1956), *J. Biol. Chem. 221*, 151.

Hood, L. (1966), *J. Cell. Physiol. 67*, 60.

Hood, L., Gray, W. R., Sanders, B. S., and Dreyer, W. J. (1967), *Cold Spring Harbor Symp. Quant. Biol. 32*, 133.

Hubbard, R. (1965), *Biochem. Biophys. Res. Commun. 19*, 679.

Katz, A. M., Dreyer, W. J., and Anfinsen, C. B. (1959), *J. Biol. Chem. 234*, 2897.

Marchalonis, J., and Edelman, G. M. (1966), *Science 154*, 1567.

Margoliash, E. (1963), *Proc. Natl. Acad. Sci. U. S. 50*, 672.

Metzger, H. (1967), *Proc. Natl. Acad. Sci. U. S. 57*, 1490.

Pan, S. C., and Dutcher, J. D. (1956), *Anal. Chem. 28*, 836.

Porter, R. R. (1967), *in* Nobel Symposium, No. 3, γ-Globulins, Killander, J., Ed., New York, N. Y., Interscience, p 81.

Press, E. M., and Piggot, P. J. (1967), *Cold Spring Harbor Symp. Quant. Biol. 32*, 45.

Press, E. M., Piggot, P. J., and Porter, R. R. (1966), *Biochem. J. 99*, 356.

Putnam, F. W. (1967), *in* Nobel Symposium, No. 3, γ-Globulins, Killander, J., Ed., New York, N. Y., Interscience, p 45.

Rüde, E., and Givol, D. (1968), *Biochem. J. 107*, 449.

Sanger, F., and Tuppy, H. (1951), *Biochem. J. 49*, 463.

Singer, S. J., and Doolittle, R. F. (1966), *Science 153*, 13.

Small, P. A., and Lamm, M. E. (1966), *Biochemistry 5*, 259.

Stark, G. R. (1967), *Methods Enzymol. 11*, 125.

Whitney, P. L., and Tanford, C. (1965), *Proc. Natl. Acad. Sci. U. S. 53*, 524.

Wilkinson, J. M., Press, E. M., and Porter, R. R. (1966), *Biochem. J. 100*, 303.