Structure of Antibodies with Shared Idiotypy: The Complete Sequence of the Heavy Chain Variable Regions of Two Immunoglobulin M Anti-Gamma Globulins

(amino-acid sequence/cross-idiotypic specificity/antibody combining site)

J. DONALD CAPRA AND J. MICHAEL KEHOE

Department of Microbiology; Mount Sinai School of Medicine, 10 East 102nd Street, New York, N.Y. 10029

Communicated by Henry G. Kunkel, June 24, 1974

ABSTRACT The complete amino acid sequence of the heavy chain variable regions of two different molecules of immunoglobulin M anti-gamma globulin has been determined. These proteins, from different human patients, had independently been shown to share idiotypic specificity. Only eight sequence differences were discernible for the entire length of their heavy chain variable regions. Five of the differences occurred outside hypervariable regions, while three were placeable within such regions. A comparison of these molecules of anti-gamma globulin with the seven human V_HIII variable region sequences presently available for immunoglobulins without known antibody activity showed that the great majority of sequence differences between the two idiotypically similar antibodies and these seven proteins were confined to hypervariable regions.

This study illustrates in precise terms a convergence of the distinct immunological properties of idiotypy, hypervariable region structure, and combining site specificity as they relate to the variable region of the immunoglobulin molecule. To a great degree these properties now appear to be a reflection of the same structural attributes of the variable region.

Subsequent to the discovery of the antigenic individuality of myeloma proteins by Slater, Ward, and Kunkel (1), a considerable body of experimental evidence has suggested that these unique characteristics can be localized to the variable regions of immunoglobulin molecules. When the concept was extended to specific antibodies (2, 3), the term "idiotypy" was introduced. In recent years, these determinants have been utilized in a wide variety of immunological studies, especially as markers for particular variable regions of immunoglobulin molecules. Recent observations concerning the inheritance of idiotypy may have profound implications for theories of antibody diversity (4).

The first evidence that the idiotypic determinants might relate directly to the antibody-combining site came from an investigation of immunoglobulin M (IgM) cold agglutinins by Williams, Kunkel, and Capra (5). This study demonstrated that immunoglobulins with similar specificities shared antigenic determinants (termed "cross-idiotypic specificity" or "shared idiotypy"), and it was postulated that the structures these proteins shared included the antibody-combining site itself. Kunkel *et al.* (6) extended these studies to the IgM antibodies against gamma globulin and showed that these proteins could be classified into two cross-idiotypic groups. Again, evidence was presented which suggested that the antibody-combining site was involved in the idiotypic crossreaction. Two proteins were selected from one of these cross-idiotypic groups for the determination of the complete amino-acid sequence of the variable regions of their heavy and light polypeptide chains. We hoped to at least approach an answer to the following questions: (1) Would two proteins with similar combining specificities and shared idiotypic specificities, selected out of a population of antibodies from genetically dissimilar individuals, share structural similarities at the level of primary structure? and (2) If there were similarities, exactly where within the variable region would they be located?

MATERIALS AND METHODS

Protein Isolation. The IgM anti-gamma globulins used in this study were first precipitated from plasma at 4°. The cryoglobulin was then washed five times with cold phosphate-buffered saline (pH 7.8) and subsequently warmed to 37°. Any material that remained as a precipitate after centrifugation at this temperature was discarded, and the supernate was applied to an A-50 DEAE-Sephadex column $(2.5 \times 100 \text{ cm})$ equilibrated with 0.1 M phosphate buffer (pH 8.2) in a 37° walk-in incubator. Under these conditions, the IgG portion of the IgM-IgG "mixed cryoglobulin" is eluted from the column. The same buffer, made 1 M with NaCl, was then applied to the column and the IgM was eluted. In most cases this material was pure by ultracentrifugal and immunochemical analyses. Occasionally proteins had to be recycled on the same column. Heavy and light chains were separated by standard procedures (7).

Protein Selection. The specific techniques used in the selection of these proteins have been described and will only be briefly reviewed here. About 40 IgM-IgG "mixed cryoglobulins" have been studied and divided into two groups on the basis of their cross-idiotypic specificities. Some proteins cannot be presently classified into either group. About 20% of IgM anti-gamma globulins have been classified in the Po group (6). This group is named for protein Pompa (abbreviated Pom in the present report). This particular group of anti-gamma globulins is characterized in a hemagglutination inhibition system, which includes erythrocytes coated with protein Pom and an antiserum made to protein Lay and absorbed with normal human serum plus a pool or IgM proteins without antibody activity against gamma globulin. Certain anti-gamma globulins inhibit this system, while others do not. Those that do are considered members of the Po group. (See Table IV, ref. 6). Despite the extensive crossreactions between these two proteins, they still contain easily detectable, individually specific

Abbreviation: IgM, immunoglobulin M.



FIG. 1. Elution profile of the cyanogen bromide digestion of the heavy chain of protein Lay on Sephadex G-100. Absorbance was monitored at 280 nm. In all gel filtration procedures, 5 M guanidine HCl was used.

antigenic determinants, and when an antiserum made against protein Lay is absorbed with protein Pom it will still react with protein Lay. The converse experiment identifies individually specific determinants on protein Pom. Thus, while these two proteins have extensive idiotypic sharing, which is more extensive than proteins of the normal pool, each still contains unique determinants.

Specificity Studies on these proteins have been reported (8). In addition, they have been performed extensively in other laboratories (9, 10). The antibody activity for both proteins has been localized to the Fab fragments of the IgM molecules. The antigenic determinants against which this activity is directed are localized to the Fc region of the IgG molecule. Extensive studies on their reactivity with immunoglobulins from the various subclasses of human IgG and various animal IgGs indicated that, while the two proteins share many specificities, there are some clear differences. In particular, protein Lay reacts with aggregated human IgG3 myeloma proteins while protein Pom does not.

Fragment Preparation. The partially reduced heavy chains were subjected to cyanogen bromide digestion (11), and the resulting fragments were fractionated on G-100 Sephadex $(5 \times 100 \text{ cm})$ in 5 M guanidine HCl. The elution diagram is shown in Fig. 1. The pattern obtained was essentially identical for both proteins, and so only protein Lay is illustrated. CBII, after further purification on a 2.5×100 -cm G-100 Sephadex column in 5 M guanidine HCl, was completely reduced and alkylated, and the resulting material fractionated as shown in Fig. 2. CBIIA was fragment 86-204 [fragments F-3, 4, and 5 of the Ou sequence of Putnam et al. (12)]. The two methionines in the fourth hypervariable region of protein Ou were not present in either protein Pom or Lay. CBIIB was fragment 1-34, which was linked to CBIIA via the Cys-Cys bond 22-98. In both proteins, fragment 34-85 was obtained relatively pure as the smallest cyanogen bromide fragment, and after further purification on G-50 Sephadex, the sequence of this fragment was determined directly on the automated sequencer. Tryptic digestion resulted in the isolation of several useful peptides that helped establish the sequence. The sequences of both intact heavy chains were determined to position 50 from the amino terminus, and, thus, an overlap between CBIIB and CBV was readily obtained. Long sequencer runs on fragment CBIIA resulted in the establishment of the sequence from 86 into the C_HI domain. No overlap was ob-



FIG. 2. Elution profile of completely reduced and alkylated Lay CBII (see Fig. 1) after a further gel filtration step. Fraction CBIIA was found to contain peptide fragment 86-204, and CBIIB was the NH₂-terminal 34-residue fragment.

tained between fragment CBV and CBIIA. Position 56 in protein Pom is reported as Asx, while it is clearly Asn in protein Lay. On two separate determinations, about 50% Asn and 50% Asp was detected at this position in protein Pom. It is likely that this is due to deamination during isolation, and in all comparisons, position 56 will be considered an Asn in both proteins.

Amino-Acid Sequence Analysis. The techniques of automated sequencing used in this laboratory have been extensively reported (13-16). An updated Beckman 890A sequencer was used for all studies. The latter parts of the study were done with newer procedures such as the fluorescamine detection system for peptides (16, 17), and the use of 4-sulfophenylisothiocyanate on lysine peptides (16, 18, 19). The sequences of some fragments of protein Pom were determined with a solid phase sequencer.

RESULTS AND DISCUSSION

The complete amino acid sequences of the variable region of the heavy chains of proteins Lay and Pom are displayed in Fig. 3. Differences in sequence have been shaded and the heavy chain hypervariable regions (16) have been enclosed. There are only eight amino-acid differences between these two heavy chain variable regions, three within hypervariable regions and five in the "framework" residues.

Fig. 4 depicts the seven presently completed heavy chain variable region sequences from human myeloma proteins of the V_HIII subgroup without known antibody activity (16, 20, 21). Proteins Lay and Pom are again displayed in this format. A line indicates identity with protein Tie. The structural differences among the seven V_HIII myelomas whose sequences were previously determined are shown in Table 1. These seven proteins differ in about 33 of the 124 positions of the $V_{\rm H}$ region, with a range of 28 to 45 differences. As can be seen from Table 1, however, the overwhelming number of differences are confined to the hypervariable regions. Thus, the typical $V_{\rm H}III$ myeloma differs from another by only six residues (range, 3-13) in the framework residues, while they differ in 27 positions (range, 25–32) in the hypervariable positions. The extreme in both situations is protein Gal, which is considerably different from the other six proteins. Nonetheless, the presence of only three amino-acid differences in all four hypervariable regions makes the similarity of proteins Pom and Lay unique among human proteins whose sequences have been determined. Two entire hypervariable regions (residues 51-68 and 101-110) are absolutely identical.



FIG. 3. The amino-acid sequence of the variable regions of the heavy chains of proteins Lay and Pom. The differences between the two sequences have been shaded and the hypervariable regions have been enclosed. Hyphens between residues have been omitted.

A considerable body of experimental evidence suggests that the major idiotypic determinants of the immunoglobulin molecule are related to the antibody-combining site. Brient and Nisonoff (22) showed that the reaction between idiotypic antibodies to antibodies directed against p-azobenzoate and p-azobenzoate was inhibited by the homologous hapten by as much as 69%. This finding suggests that the region of the combining site for this particular hapten is, in fact, a major idiotypic determinant. It is of course possible that combination of the hapten with the site results in a conformational change in other portions of the variable region that alters the idiotypic determinants. Similar results have been obtained by Sher and Cohn (23), who showed that phosphorylcholine inhibited the reaction between a mouse antibody against phosphorylcholine myeloma protein and its anti-idiotypic antiserum. More recent studies linking the cross-idiotypic determinants to the combining site comes from the work of Weigert et al. (24) with antibodies against α (1 \rightarrow 6) dextran mouse myelomas. These authors introduced the term "ligandmodifiable idiotype" to indicate an idiotypic reaction that is inhibited by specific ligand.

Conclusions concerning the relationship between the hypervariable regions and the antibody-combining site come primarily from studies with affinity labels. In light chains, affinity labels have been localized near or within hypervariable regions (25–27), providing direct support for the general concept that hypervariable regions participate directly in the antibody-combining site. Most of the recent work on the heavy chain supports this general view. For example, Ray and Cebra localized affinity labels to the first (31–37) and fourth (101–110) heavy chain hypervariable regions (28); Haimovich *et al.* (29) localized an affinity label to residue 54 of mouse myeloma protein 315; and Press and coworkers have localized affinity labels at or near the fourth hypervariable region in rabbit antibodies (30). Koo and Cebra have recently affinity labeled a lysyl residue at position 59 of the heavy chain of guinea pig antibodies against *p*-azobenzene arsonate (31). Thus, the general idea has emerged that the same regions of the molecule that show the highest degree of sequence variation are near or part of those exact regions of the heavy chain where affinity labels have been localized.

Another line of evidence linking the hypervariable regions to the antibody-combining site comes from sequence studies on pooled immunoglobulins from several species. These studies indicate that no definitive amino acid sequence can be obtained in most hypervariable positions in pools (32, 33), while specifically purified antibody does have a defined sequence throughout at least the first three hypervariable regions (33). Finally, crystallographic studies on human (34, 35) and mouse (36) myeloma proteins with antibody activity indicate that, in general, the hypervariable regions line a "pocket" or "groove," which is thought to represent the combining site itself.

TABLE 1. Variable region structural differences among $V_{H}III$ myelomas and two IgM proteins with
cross-idiotypic specificity

	Total residues (124)	Framework residues (83)	Hypervariable residues (41)
V _H III myelomas (Tie,			
Was, Jon, Zap, Tur,	33	6	27
Gal, Nie)	(28–45)	(3–13)	(25–32)
Pom/Lay IgM anti- gamma globulins	8	5	3



FIG. 4. The seven complete human $V_{\rm H}$ III sequences from proteins without known antibody activity compared with proteins Lay and Pom, two IgM anti-gamma globulins with shared idiotypy. Identities with protein Tie are indicated by a line. Hyphens between residues have been omitted.

The genetic origin of hypervariable regions has not been established. The clear demonstration of the inheritance of idiotypic determinants (37–39), and the indication from hapten inhibition studies (22-24) and the present study that the hypervariable regions are the major idiotypic determinants, could very readily be explained if the information encoding the hypervariable region sequences is carried in the germ line. This would either vastly expand the number of germ line genes so that even the most stringent somatic mutation theories would become in effect "multi-gene," or it would indicate that there are, in fact, vast numbers of germ line genes and that somatic mutation plays little if any role in the generation of antibody diversity.

The complete identity of two of the four hypervariable regions in these two anti-gamma globulins from genetically distinct individuals, however, suggests a third possibility: gene interaction (J. D. Capra and T. J. Kindt, Immunogenetics, in press). According to this view, the hypervariable region sequence would be carried on a separate piece of DNA, and the information contained within this DNA region would be combined with the information for the rest of the variable region to form a mRNA molecule encoding for the entire variable region. This concept has the advantage of allowing for the more stable portions of the variable region (those portions containing the phylogenetically associated residues and the variable region allotypes in certain species) to be encoded by a single (or a very few) germ line gene, while the hypervariable region, idiotypic and combining site sequences would be carried in the germ line, selected upon by evolution, and inherited as indicated above.

Finally, it is possible that neither the germ line nor gene interaction hypotheses are correct and that the hypervariable region sequences arise by localized somatic mutation events (40). A major difficulty with this proposal is explaining how two genetically unrelated individuals can generate, by a random process such as somatic mutation, two different proteins (e.g., Lay and Pom) that have one or more hypervariable regions of absolutely identical amino acid sequence.

No matter what the actual genetic origin of hypervariable regions, the studies reported here imply that the antibodycombining site is primarily determined by such regions, localized near a groove or pocket, and that antigenically potent portions of this exposed groove are responsible for the major idiotypic determinants of immunoglobulin molecules. Thus, three independently conceived immunological concepts (idiotypes, hypervariable regions, and the combining site) should be considered as an interrelated ensemble in any further hypotheses concerning the nature of antibody diversity.

Note Added in Proof. Recently, Lifter *et al.* (1974) [Biochemistry 13, 3567–3571] have shown that an affinity label can be localized to positions 33 and 88 of the heavy chain of the murine myeloma protein MOPC 460, which possesses anti-dinitrophdnol activity. Thus, all four heavy chain hypervariable regions have now been affinity labeled in one system or another.

We thank Drs. Henry Metzger, Marvin Stone, Ralph Schrohenloher, and Stephen Ritzman for providing plasma from patients Lay and Pom. Vickie Ecker, Bonnie Gerber, Ellen Bogner, and Donna Atherton rendered invaluable technical assistance. We also thank Dr. Henry Kunkel, who provided advice and encouragement during all phases of this study. This work was supported in part by grants from the National Science Foundation (GB 17046), and the U.S. Public Health Service (AI 09810), and a Grant-in-Aid from the New York Heart Association. J.D.C. is the recipient of NIH Career Development Award 6-K4-GM-35, 190, and J.M.K. is an Established Investigator of the American Heart Association.

- Slater, R. J., Ward, S. M. & Kunkel, H. G. (1955) J. Exp. Med. 101, 85-108.
- Kunkel, H. G., Mannik, M. & Williams, R. C. (1963) Science 140, 1218–1219.
- 3. Oudin, J. (1966) Proc. Roy. Soc. Ser. B 166, 207-219.
- 4. Capra, J. D. & Kehoe, J. M. (1974) Scand. J. Immunol. 3, 1-4.
- Williams, R. C., Jr., Kunkel, H. G. & Capra, J. D. (1968) Science 161, 379–381.
- Kunkel, H. G., Agnello, V., Joslin, F. G., Winchester, R. J. & Capra, J. D. (1973) J. Exp. Med. 137, 331-342.
- Fleischman, J. G., Porter, R. R. & Press, E. M. (1963) Biochem. J. 88, 220–228.
- Capra, J. D., Kehoe, J. M., Winchester, R. J. & Kunkel, H. G. (1971) Ann. N.Y. Acad. Sci. 190, 371-381.
- 9. Stone, M. J. & Metzger, H. (1967) Cold Spring Harbor Symp. on Quant. Biol. 32, 83.
- 10. Schrohenloher, R. (1973) J. Immunol. 111, 1699-1711.
- Gross, E. & Witkop, X. (1962) J. Biol. Chem. 237, 1856– 1863.
- Putnam, F. W., Florent, G., Paul, C., Shinoda, R. & Shimizu, A. (1973) Science 182, 287-291.
- 13. Capra, J. D. (1971) Nature New Biol. 230, 61-63.
- Kehoe, J. M. & Capra, J. D. (1971) Proc. Nat. Acad. Sci. USA 68, 2019-2021.
- Capra, J. D., Wasserman, R. W. & Kehoe, J. M. (1973) J. Exp. Med. 138, 410-427.
- Capra, J. D. & Kehoe, J. M. (1974) Proc. Nat. Acad. Sci. USA 71, 845–848.
- 17. Udenfriend, S., Stein, S., Bohlen, P., Dairman, W., Leimgruber, W. & Wiegele, M. (1972) Science 178, 881-882.
- Inman, J. K., Hannon, J. E. & Appella, E. (1972) Biochem. Biophys. Res. Commun. 46, 2075–2081.
- Foster, J. A., Bruenger, E., Gray, W. R. & Sandberg, L. B. (1973) J. Biol. Chem. 248, 2876-2879.
- Ponstingl, H., Schwarz, J., Reichel, W. & Hilschman, N. (1970) Hoppe-Seyler's Z. Physiol. Chem. 351, 1591–1594.
- Watanabe, S., Barnikol, H. U., Horn, J., Bertram, J. & Hilschman, N. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1505-1509.
- Brient, B. W. & Nisonoff, A. (1970) J. Exp. Med. 132, 951– 962.
- 23. Sher, A. & Cohn, M. (1972) Eur. J. Immunol. 2, 319-326.
- Weigert, M., Raschke, W. C., Carson, D. & Cohn, M. (1974) J. Exp. Med. 139, 137-147.
- Goetzl, E. J. & Metzger, H. (1970) Biochemistry 9, 1267– 1278.
- 26. Franek, F. (1971) Eur. J. Biochem. 19, 176-183.
- 27. Chesebro, B. & Metzger, H. (1972) Biochemistry 11, 766-771.
- 28. Ray, A. & Cebra, J. J. (1972) Biochemistry 11, 3647-3657.
- Haimovich, J., Eisen, H. N., Hurwitz, E. & Givol, D. (1972) Biochemistry 11, 2389-2397.
- Press, E. M., Fleet, G. W. J. & Fisher, C. E. (1971) in Progress in Immunology, ed. Amos, B. (Academic Press, New York), p. 233.
- 31. Koo, P. H. & Cebra, J. J. (1974) Biochemistry 13, 184-195.
- Cebra, J. J., Givol, D. & Porter, R. R. (1968) Biochem. J. 107, 69-70.
- Cebra, J. J., Ray, A., Benjamin, D. & Birshtein, B. (1971) in *Progress in Immunology* (Academic Press, New York), p. 269.
- Poljak, R. J., Amzel, L. M., Avery, H. P., Chen, B. L., Phizackerley, R. P. & Saul, F. (1973) Proc. Nat. Acad. Sci. USA 70, 3305-3310.
- Amzel, L. M., Poljak, R. J., Saul, F., Varga, J. M. & Richards, F. F. (1974) Proc. Nat. Acad. Sci. USA 71, 1427–1430.
- Padlan, E. A., Segal, D. M., Spande, T. F., Davies, D. R., Rudikoff, S. & Potter, M. (1973) Nature New Biol. 245, 165-167.
- Eichmann, K. & Kindt, T. J. (1971) J. Exp. Med. 134, 532-552.
- Pawlak, L. L., Mushinski, E. B., Nisonoff, A. & Potter, M. (1973) J. Exp. Med. 137, 22-31.
- Eichmann, K. & Berek, C. (1973) Eur. J. Immunol. 3, 599-601.
- 40. Cohn, M. (1971) Ann. N.Y. Acad. Sci. 190, 529-584.