Proc. Natl. Acad. Sci. USA
Vol. 73, No. 12, pp. 4471–4473, December 1976
Biochemistry

Some correlations between specificity and sequence of the first complementarity-determining segments of human kappa light chains

(hypervariable regions/antibody specificity)

ELVIN A. KABAT*, TAI TE WU†, AND HOWARD BILOFSKY‡

*Departments of Microbiology, Human Genetics and Development, and Neurology, Columbia University, and the Neurological Institute, Presbyterian Hospital, New York, N.Y. 10032 and National Cancer Institute, Bethesda, Maryland 20014; †Departments of Engineering Sciences, and Biochemistry and Molecular Biology, Northwestern University, Evanston, Illinois 60201; and ‡Bolt Beranek and Newman, Inc., Cambridge, Massachusetts 02138

Contributed by Elvin A. Kabat, October 6, 1976

ABSTRACT **Examination of the sequences of the first complementarity-determining segments of the light chains of two IgM cold agglutinins against blood group I, four monoclonal IgM antibodies against IgG proteins, and of three Bence Jones proteins provides clues for predicting which residues contribute to antibody specificity and indicates that these predictions may be tested by evaluating recovery of antibody activity and specificity when various light chains are recombined with homologous and heterologous heavy chains.**

The verification, by x-ray crystallographic studies of Fab′ fragments (1, 2), Bence Jones proteins (3, 4), and Fv fragments (5, 6) (see also ref. 7) and by affinity labeling (8), of the prediction (9, 10) that the hypervariable regions are the complementarity-determining segments and form the antibody-combining site has substantially simplified the problem of understanding the nature of immunological specificity (11). Thus, the noncomplementarity-determining segments of the light chain of the variable region (V_L) and the heavy chain of the variable region (V_H) form a three-dimensional framework of twisted β-sheets so oriented spatially as to position the complementarity-determining segments at the tip of the molecule available for reacting with the antigenic determinant (1–7). It becomes possible therefore to correlate antibody specificity with sequence data on the complementarity-determining segments as such data become available and also to predict structures of other combining sites by inserting sequences of complementarity-determining segments when these have been determined for antibodies of known specificities. Thus, using the structure of the Fab′ fragment of the phosphocholine-binding myeloma protein McPC603 (2, 7), a model was constructed of a combining site for an antibody to the type III capsular polysaccharide of the pneumococcus (12) and for an antibody site complementary to 2,4-dinitrophenyl (13); the validity of these models may ultimately be decided by x-ray crystallography.

As more sequence data become available, the role of individual residues in antibody-combining sites may be evaluated. By use of a data bank (14) of variable region sequences aligned for maximum homology and stored in the PROPHET computer system (15), a statistical analysis of the occurrence in complementarity-determining segments of pairs of residues more frequently than expected by chance permitted the recognition of two contacting residues Tyr 33 and Glu 35 in the first complementarity-determining segment of the heavy chain of phosphocholine-binding mouse myeloma proteins, the assignment of a structural role to Met 34, and the recognition of the probable contribution of Phe 32 to the orientation of Tyr 33 (16). The sequence Phe-Tyr-Met-Glu has not been found in any other protein tabulated (14, 17–19) and has been seen only in phosphocholine-binding mouse myeloma proteins and only at positions 32 to 35 of the H chain (14). A human phosphocholine-binding myeloma (20) had the sequence Phe-Tyr-Met-Asp at positions 32 to 35 (21).

As more sequences of complementarity-determining segments are accumulated, similarities and differences may provide insights bearing on antibody specificity. While cross-idiotypic specificities (22, 23) provide important clues to structural similarities in complementarity-determining segments, the lack of insight into what an idiotypic determinant is in terms of sequence seriously hampers interpretations.

Identical first complementarity-determining segments in a $V_\lambda II$ and a $V_\lambda V$ human light chain (24), despite 21 amino acid differences elsewhere in their variable regions, has been described. In addition, a $V_\kappa I$ and a $V_\kappa III$[§], light chains of two IgM with anti-IgG specificity and cross-idiotypic specificity were found to have identical third complementarity-determining segments (25). The finding that the heavy chain of a monoclonal IgM with antibody specificity against blood group I yielded an anti-idiotopic serum that reacted equally well with the isolated light and heavy chains (26), if ultimately supported by identical sequences in one or more complementarity-determining segments of each of these chains, would contribute substantially to the understanding of the structural basis of idiotypic specificity.

From the data bank we have examined the first complementarity-determining segments of human V_κ chains with respect to their antibody specificities.

Table 1 lists the data obtained and their sources. Three IgM agglutinins, DRE, a cold agglutinin with anti-blood group I activity, and two non-cold agglutinins, GLO and SIE, with anti-gamma globulin activity, have identical first complementarity-determining segments of 12 residues. The two anti-gamma globulins show cross-idiotypic specificity belonging to the Wa group (32). The anti-blood group I agglutinin is known to be specific for oligosaccharide structures involving D-galactose and *N*-acetyl-D-glucosamine (33 34); the specificities of the antigamma globulins are not known, but are not directed toward the carbohydrate determinants since they still react with IgG treated to destroy carbohydrate (H. G. Kunkel, personal communication). Moreover, protein LAY reacted with IgG but did not react with IgA, both of which contain similarly linked oligosaccharide chains (35).

Two additional proteins, IKE and TAK, differ from DRE, GLO, and SIE at position 30, assuming position 27 of TAK to be Gln (footnote[a], Table 1). IKE, a Bence Jones protein without known antibody specificity, has Asn instead of Ser, while in TAK, another IgM cold agglutinin with anti-I specificity, the residue has not been identified but is either Gly or Trp (28).

Abbreviations: V_L and V_H, variable regions of the light and heavy chain, respectively.

§ Called $V_\lambda II$ in ref. 25.

Table 1. Sequences of the first hypervariable region of human V_κ III light chains[a]

Protein	Residue 24	25	26	27	27a	28	29	30	31	32	33	34	Antibody activity	Ref.
DRE	Arg	Ala	Ser	Gln	Ser	Val	Ser	Ser	Ser	Tyr	Leu	Ala	Anti-blood group I; IgM cold agglutinin	27, 28
GLO													Anti-human IgG; IgM Wa idiotype	23
SIE													Anti-human IgG; IgM Wa idiotype	23
TAK				Glx[b]				?[c]					Anti-blood group I; IgM cold agglutinin	28
IKE								Asn					Bence Jones protein; no known antibody activity	23
TI								Asn		Phe			Bence Jones protein; no known antibody activity	29
WIL								Asn		Met			Bence Jones protein; no known antibody activity	31
POM[d]						Ile		Asn					Anti-IgG1; IgM Po idiotype	25
V_κI LAY	Gln				Asn		Asn	Ala	( )[e]			Asn	Anti-human IgG1 and anti-human IgG3; IgM Po idiotype	25, 31

[a] All residues are the same as those in DRE except as indicated.
[b] The sequences of 22 V_κIII proteins have been determined: 16 have Gln, 5 have Glx, and 1 has Ser. Capra *et al.* (27) have Gln for residue 27; Gergely *et al.* (28) determined it only as Glx.
[c] Gly or Trp are considered to be the only possibilities (28).
[d] The sequence originally reported (31) has been revised.
[e] Deletion.

Three proteins TI, WIL, and POM differ from DRE, GLO, and SIE by two amino acid residues. The first two are Bence Jones proteins without known antibody activity having Asn at position 30 like IKE but having Phe and Met, respectively, at position 32. The third POM, an IgM with anti-IgG1 specificity (31), differs in idiotypic specificity from GLO and SIE and has Ile at position 28 instead of Val and Asn at position 30.

The last protein, LAY, has a V_κI light chain, belongs to the Po group in idiotypic specificity, and is an IgM but differs from POM in showing both anti-IgG1 and anti-IgG3 specificity (31, 35).

These data permit certain testable predictions to be made. (*i*) The complete sequences of the variable regions of both chains of LAY and POM, as established by Klapper and Capra (25) have shown that the heavy chains each have identical second and third complementarity-determining segments and that the light chains are identical in their third and in residues 50 through 54 of their second complementarity-determining segment but differ at positions 55 and 56, which were considered (9) as part of the second complementarity-determining segment. LAY and POM differ by only one residue in the first complementarity-determining segment of their heavy chains having Ala and Ser, respectively, at position 31; the first complementarity-determining segments of their light chains differ at six positions (Table 1) and, in addition, that of LAY is one residue shorter.

Since these are the only differences in complementarity-determining segments, they must be responsible for the difference in specificity, LAY reacting equally well with IgG1 and IgG3 while POM is more specific for IgG1. Moreover, since their heavy chains differ by only one residue, the bulk of the specificity difference probably resides in their light chains. Chain recombinations with these two proteins to evaluate the effects of antibody specificity would be of interest.

(*ii*) The three Bence Jones proteins IKE, TI, and WIL most probably do not come from Waldenström IgM with anti-IgG activity or from cold agglutinins with anti-I specificity since light chains from such proteins almost never occur in urine as Bence Jones proteins. Thus the finding that TI and WIL differ at positions 28 and 32 and that IKE differs only at position 28 from POM implicates the Ile at position 28 as an important contacting residue of the anti-IgG1 site, with this position and the substitutions at other positions in LAY making for the differences in specificity. The effect of recombining the heavy chains of LAY and POM with the three Bence Jones proteins on antibody and idiotypic specificity would also contribute insights into the role of various residues on specificity.

These sequences may also be correlated with the x-ray structure of protein REI, a V_κI Bence Jones protein dimer with one less residue in the first complementarity-determining segment than all of the chains in Table 1, the extra residue being listed as 27a. Examination of the REI structure (5) indicates that residues 24, 25, 26, and 33 cannot be contacting residues for antigen in the combining site; the side chain of 24 points in the opposite direction, residues 25 and 26 are shielded from the site by other residues, and residue 33 points toward the interior of the molecule. In the second complementarity-determining segment, residues 54 and 56 also cannot be contacting antigens because the side chain of residue 54 points in the opposite direction and residue 56 is at some distance from the site. Thus, based on the REI model, the binding residues of POM could be Gln 27, Ile 28, Ser 29, Ser 30, Ser 31, Tyr 32 and Ala 34, and Ala 55. The extra residue at 27a might cause rearrangement of some adjacent residues but it is noteworthy that, in addition to residue 27a, LAY differs from POM at positions 28, 29, 30, 34, and 55. Thus, these positions plus position 31 in the heavy chain must be responsible for the difference binding specificity for the IgG subgroups.

The light chains of the two IgM proteins GLO and SIE with anti-IgG specificity are identical in first complementarity-determining segments to the IgM cold agglutinin DRE with anti-I specificity and differ only at position 30 from the second anti-I protein TAK. It is unlikely that anti-I specificity, which in one instance (34) has been shown to recognize at least the DGalβ1 → 4DGlcNAcβ1 → 6 structure, involves predominantly the first complementarity-determining segment, since this

segment is implicated in the specificity differences between LAY and POM. Although complete sequences are not available, recombination experiments (see ref. 36) with the heavy chains of the anti-I proteins and the various light chains might also contribute to defining the contribution of the various complementarity-determining segments to blood group I specificity.

We are indebted to Dr. Otto Epp for describing the location of the residues in the REI model (5). This work is sponsored by the National Cancer Institute, National Institute of Allergy and Infectious Diseases, National Institute of Arthritis, Metabolism and Digestive Diseases, National Institute of General Medical Sciences, Division of Research Resources (Contract no. N01-RR-4-2147), National Institutes of Health. T.T.W. is a Career Research Development Awardee (5-K01-AI-70497), National Institutes of Health. The work was aided in part by a grant from the National Institutes of Health, 1-R01-GM21482-01, to T.T.W. and by a grant from the National Science Foundation, BMS-72-02219-A04 to E.A.K.

1. Poljak, R. J., Amzel, L. M., Chen, B. L., Phizackerley, R. P. & Saul, F. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 3440–3444.
2. Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M. & Davies, D. R. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4298–4302.
3. Schiffer, M., Girling, R. L., Ely, K. R. & Edmundson, A. B. (1973) *Biochemistry* **12,** 4620–4631.
4. Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, M. & Panagiotopoulos, N. (1975) *Biochemistry* **14,** 3953–3961.
5. Epp, O., Colman, P. M., Fehlhammer, H., Bode, W., Schiffer, M. & Huber, R. (1974) *Eur. J. Biochem.* **45,** 513–524.
6. Fehlhammer, H., Schiffer, M., Epp, O., Colman, P. M., Lattman, E. E., Schwager, P. & Steigemann, W. (1975) *Biophys. Struct. Mech.* **1,** 139–146.
7. Davies, D. R., Padlan, E. A. & Segal, D. M. (1975) *Annu. Rev. Biochem.* **44,** 639–667.
8. Givol, D. (1974) *Essays Biochem.* **10,** 73–103.
9. Wu, T. T. & Kabat, E. A. (1970) *J. Exp. Med.* **132,** 211–250.
10. Kabat, E. A. & Wu, T. T. (1971) *Ann. N.Y. Acad. Sci.* **190,** 382–392.
11. Kabat, E. A. (1976) *Structural Concepts in Immunology and Immunochemistry* (Holt, Rinehart and Winston, New York), 2nd ed.
12. Davies, D. R. & Padlan, E. A. (1975) *Royal Society of Medicine Symposium, Rockefeller University, Oct. 20–22, 1975* (Raven Press, New York), in press.
13. Padlan, E. A., Davies, D. R., Pecht, I., Givol, D. & Wright, C. (1976) *Cold Spring Harbor Symposium on Origins of Lymphocyte Diversity,* in press.
14. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1976) *Variable Regions of Immunoglobulin Chains. Tabulations and Analyses of Amino Acid Sequences* (Bolt Beranek and Newman, Cambridge, Mass.)
15. Raub, W. F. (1974) *Fed. Proc.* **33,** 2390–2392.
16. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 617–619.
17. Dayhoff, M. E. (1972) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Md.), Vol. 5.
18. Dayhoff, M. E. (1973) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Md.), Suppl. I.
19. Dayhoff, M. E. (1976) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Md.), Suppl II.
20. Riesen, W., Rudikoff, S., Oriol, R. & Potter, M. (1975) *Biochemistry* **14,** 1052–1057.
21. Riesen, W. F., Braun, D. G. & Jaton, J. C. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2096–2100.
22. Williams, R. C., Kunkel, H. G. & Capra, J. D. (1968) *Science* **161,** 379–381.
23. Capra, J. D. & Kehoe, J. M. (1975) "Hypervariable regions, idiotypy and the antibody-combining site." *Adv. Immunol.* **20,** 1–40.
24. Wu, T. T., Kabat, E. A. & Bilofsky, H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 5107–5110.
25. Klapper, D. G. & Capra, J. D. (1976) *Ann. Immunol. (Paris)* **C127,** 261–272.
26. Kobzik, L., Brown, M. C. & Cooper, A. G. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 1702–1706.
27. Capra, J. D., Kehoe, J. M., Williams, R. C., Jr., Feizi, T. & Kunkel, H. G. (1972) *Proc Natl. Acad. Sci. USA* **69,** 40–43.
28. Gergely, J., Wang, A. C. & Fudenberg, H. H. (1973) *Vox Sang.* **24,** 432–440.
29. Suter, L., Barnikol, H. U., Watanabe, S. & Hilschmann, N. (1969) *Z. Physiol. Chem.* **340,** 275–278.
30. Suter, L., Barnikol, H. U., Watanabe, S. & Hilschmann, N. (1972) *Z. Physiol. Chem.* **353,** 189–208.
31. Capra, J. D., Kehoe, J. M., Winchester, R. J. & Kunkel, H. G. (1971) *Ann. N.Y. Acad. Sci.* **190,** 371–380.
32. Kunkel, H. G., Agnello, V., Joslin, F. G., Winchester, R. J. & Capra, J. D. (1973) *J. Exp. Med.* **137,** 331–342.
33. Marcus, D. M., Kabat, E. A. & Rosenfield, R. E. (1963) *J. Exp. Med.* **118,** 175–194.
34. Feizi, T., Kabat, E. A., Vicari, G., Anderson, B. & Marsh, W. L. (1971) *J. Immunol.* **106,** 1578–1592.
35. Metzger, H. (1967) *Proc. Natl. Acad. Sci. USA* **57,** 1490–1497.
36. Manjula, B. H., Glaudemans, C. P. J., Mushinski, E. G. & Potter, M. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 932–936.