0273-2289/94/4702-03—0107$03.40

From Esoteric Theory to Therapeutic Antibodies

TAI TE WU

Departments of Biochemistry, Molecular Biology and Cell Biology, of Biomedical Engineering, and of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208

ABSTRACT

Theoretical analyses of amino acid and nucleotide sequences of immunoglobulins have provided a unique approach to the understanding of structure and function of antibodies. Variability plots unambiguously identified that the antibody-combining site is formed by six short complementarity determining regions (CDRs), three each from light and heavy chains. Since three-dimensional (3-D) foldings of framework regions (FRs) are similar among different antibodies, the 3-D configurations of CDRs from a specific antibody can be predicted based on their amino acid sequences. The resulting structure forms a compact surface that fits the antigen molecule tightly. The third CDR of the heavy chain (CDRH3), which is coded by the D-minigene together with the N- and/or P-segments, appears to play a unique role in fine fitting between antibody and antigen molecules. In order to maintain biological activities, the grafting of mouse CDRs onto human FRs should closely match the human FR amino acid sequences with the original mouse antibody. Similarities between human and mouse FR sequences that have been preserved for over 20 million years of evolution can be a useful tool in humanizing mouse antibodies.

Index Entries: Antibodies; amino acid sequences; nucleotide sequences.

INTRODUCTION

In 1969, Elvin A. Kabat and I started our collaboration by analyzing available amino acid and nucleotide sequences of antibodies, in order to understand the structure and function of these important biological macromolecules. Our other collaborators in this endeavor during the past 24 years have been Howard Bilofsky, Steve Coutre, Garry Hovis, George Johnson, Priscilla Wilkins Stevens, and Janette Stanford.

Several basic facts were known about antibodies in 1969. They are protein molecules that have evolved to bind antigen molecules tightly, i.e., nearly irreversibly. Light chains of immunoglobulins from patients with multiple myeloma were excreted in their urine as Bence Jones proteins (*1*). Variable and constant regions of light chains have been identified based on their amino acid sequences (*2*). Subgroups of κ light chains were defined by their N-terminal variable region amino acid sequences (*3*). A number of positions in the variable region showed extensive amino acid substitutions, whereas several other positions were conserved (*3–6*). The most pressing problem at that time was to try to locate segments in the variable region that might bind antigen molecules (*7*).

The function of antibodies is different from protein enzymes. Enzymes can either bind substrate reversibly, e.g., hemoglobin binding oxygen molecules, or convert substrates to products, e.g., glycerol dehydrogenase converting glycerol to dihydroxyacetone. Enzymes with similar functions have similar structures, and those amino acid residues that form the active site are highly conserved. Antibodies are protein molecules with similar structures but different functions, since each one will bind a different antigen nearly irreversibly. As a result, amino acid residues in their active sites must vary extensively in order to generate different binding properties. Homologous amino acid sequences of enzymes with similar functions or antibodies with different properties can be aligned. From these aligned sequences, we tried to define a quantitative measure that might be able to identify their active sites precisely.

Although there were several statistical measures available, none of these seemed adequate to analyze the above problem. We, therefore, tried to introduce a simple quantity, known as variability (*8*), which could be calculated for each position of aligned homologous amino acid sequences. Variability is defined for each position as:

$$V(p) = \text{number of different amino acid residues / frequency of the most common one} \quad (1)$$

where V is variability, a function of position p. It can vary from 1 (for an invariant position) to 400 (for a position with all 20 different amino acid substitutions occurring with the same frequency). This esoteric theoretical number, when plotted for aligned variable region amino acid sequences of immunoglobulin light chains available in 1970, gave three distinct peaks of high variability, located at positions 24–34, 50–56, and 89–97 (*8*).

Based on the above logical argument, we thus postulated that the segments of light chains under these three peaks together with three similar segments of heavy chains with hypervariability, positions 31–35B, 50–65, and 95–102 (*9*), would form the antibody-combining site. These short segments were designated as complementarity determining regions (CDRs), and the rest of the variable region as framework regions (FRs). For comparison, a similar variability plot for amino acid sequences of cytochromes c from different species showed that the active site was located at positions with very low variability (*10*) as expected, since they all have the same function of transporting electrons.

In 1970, other researchers were skeptical about our result, primarily because of the fact that I had just published a paper in 1969 (*11*) suggesting that DNA could exist in at least two distinct forms: the double helix and a four-stranded structure consisting of two mutually intercalating straight ladders with or without a slight twist. The four sugar-phosphate backbones appeared in an up-down-up-down configuration that was also recently proposed by Gehring et al. (*12*) based on NMR studies of a short d(TCCCCC) segment.

Fortunately, within a few years, in 1973, an immunoglobulin Fab fragment (*13*) and a light-chain dimer (*14*) were crystallized, and X-ray diffraction studies showed that the six CDRs indeed form a compact surface that could bind antigen molecules. Many other immunoglobulin structures were subsequently determined and showed the same features (*see 15*). The final verification of our prediction was achieved in 1986 (*16*), i.e., 16 years after the use of our esoteric theoretical measure. Poljak's group cocrystallized lysozyme with the Fab fragment of an antilysozyme antibody with high binding affinity. In this case and other later studies (*see 15*), all six CDRs were involved in binding the antigen, and provided a perfect fit between the antibody and antigen molecules.

We also proposed in 1970 (*8*) that in order for a cell to synthesize a different antibody, one simple possibility was to replace the gene segments coding for CDRs by a different set. This process, known as gene conversion (*17*), has subsequently been discovered to be the mechanism used by chickens to make different light and heavy chains of antibodies (*18,19*). It may also play an important role in the human, where many immunoglobulin pseudogenes have recently been identified (*20*), as well as in other species. Similarly, on analyzing the amino acid sequences of FRs, many of which were shared by different light and heavy chains, we further suggested that FRs might also be coded by minigenes (*21*). The discovery of J minigenes in light chains (*22*) and the D and J minigenes in heavy chains (*23,24*), provided a wealth of genetic processes that could be used to generate antibody diversity. However, the presence of human D-minigene nucleotide segments in CDRH2 of several human and mouse heavy-chain genes (*25*) is unexpected and still unexplained.

After locating CDRs and FRs of the variable region of light and heavy chains of antibodies, we explored the possibility of whether we could

predict the 3-D structures of antibody-combining sites based on their CDR amino acid sequences. This classical problem of protein primary structure determining its 3-D folding had been studied by many researchers throughout the world (*26*). Various methods were developed to predict locations of α-helices, β-strands, and so forth. Unfortunately, most of these techniques were of little use to us, since antibody CDRs were relatively short and most likely containing neither α-helices nor significant β-strands.

Instead, we developed a method to estimate the ϕ and ψ angles of the protein backbone (*27*). For each amino acid residue, there were many theoretically allowable values for the ϕ and ψ angles. In order to reduce the ϕ and ψ choices for a specific residue at a given position in the CDR, we took into consideration the influences of nearest neighboring amino acid residues (*28*). This required the use of amino acid sequences and atomic coordinates of proteins, the 3-D structures of which had been determined by X-ray diffraction studies. Such data on these reference proteins have been stored at the Brookhaven Protein Data Bank (*29*). ϕ and ψ angles could be calculated from atomic cooordinates as follows (Fig. 1):

$$P = \cos^{-1}[(\vec{C}\,X\,\vec{N} / |\,\vec{C}\,X\,\vec{N}\,|) \cdot (\vec{N}\,X\,\vec{C}_{-1} / |\,\vec{N}\,X\,\vec{C}_{-1}\,|)]$$

$$\text{If } \vec{C}\,X\,\vec{N} \cdot \vec{C}_{-1} > 0, \text{ then } \phi = P - 180°.$$
$$< \qquad = 180° - P. \tag{2}$$

$$Q = \cos^{-1}[(\vec{C}\,X\,\vec{N} / |\,\vec{C}\,X\,\vec{N}\,|) \cdot (\vec{N}_{+1}\,X\,\vec{C} / |\,\vec{N}_{+1}\,X\,\vec{C}\,|)]$$

$$\text{If } \vec{C}\,X\,\vec{N} \cdot \vec{N}_{+1} > 0, \text{ then } \psi = 180° - Q.$$
$$< \qquad = Q - 180°. \tag{3}$$

For an amino acid residue at position n, we matched tripeptide $(n-1)$ (n) $(n+1)$, and dipeptides $(n-1)$ (n) ( ), $(n-1)$ ( ) $(n+1)$, and ( ) (n) $(n+1)$ in CDRs with those in reference proteins (*27*). The ϕ and ψ angles of the middle amino acid residues were then used to estimate possible ϕ and ψ combinations of the original residue in the CDRs. Usually, several choices were possible. Initially, we tried this method on a small peptide hormones, oxytocin, in collaboration with Levinthal's group (*30*), and the predicted 3-D structure was reasonable. We also used this method to estimate ϕ and ψ angles of horse cytochrome c (*31*), and the predicted values were generally in good agreement with experimentally obtained angles.

Since the 3-D structures of FRs were quite similar among different antibodies (*see* ref *15*), our task was basically to try to replace a set of six CDRs by different sets, identical to the idea of lock-and-key for enzymes. Starting with a known Fv or Fab 3-D structure, we could cut the CDRs off and put on a new set. Thus, each CDR had to originate at a fixed position with defined orientation and end up at a given position also with prescribed orientation. Using the estimated ϕ and ψ angles for each amino

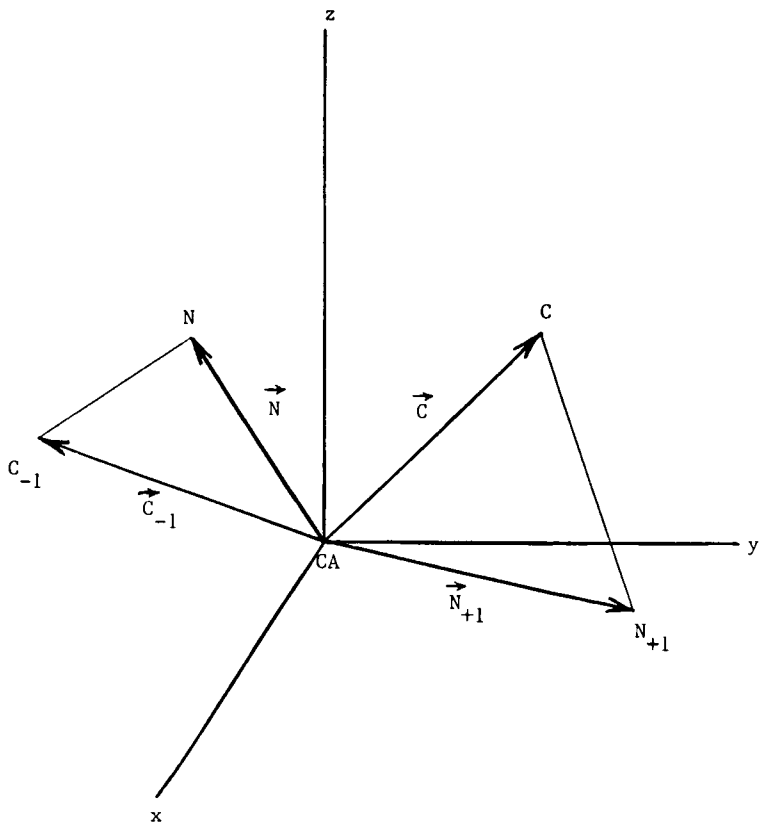

Fig. 1. To calculate ϕ and ψ angles of a specific amino acid residue, the α carbon, CA, is place at the origin of the x, y, z coordinate system. Vectors, C_{-1}, N, C, and N_{+1}, are defined as shown.

acid residue of every CDR, we started from one end and tried to reach the other end. In order to fit the required orientations, the quantity to be minimized by varying the ϕ and ψ choices within 30° was the sum of squares of distances between theoretical and actual α carbon atoms of three amino acid residues beyond the end of each CDR (32). This required a massive amount of computing time. Fortunately, we had access to the NIH-supported PROPHET computer system, which was available to several hundred user groups. We, being one user group, managed to use more than 80% of the total CPU time. We applied this method to predict the backbone conformations of the antibody-combining sites of MOPC315 (32), MOPC617 (33), and MOPC104E (34). For MOPC104E, the side-chain configurations were also predicted (34). We were hoping that this immunoglobulin would be crystallized in 1982, since its combining site was of great interest because of its ability to bind polysaccharides of different lengths. Then, our predicted structure could be compared with that determined by X-ray diffraction studies. Unfortunately, even today,

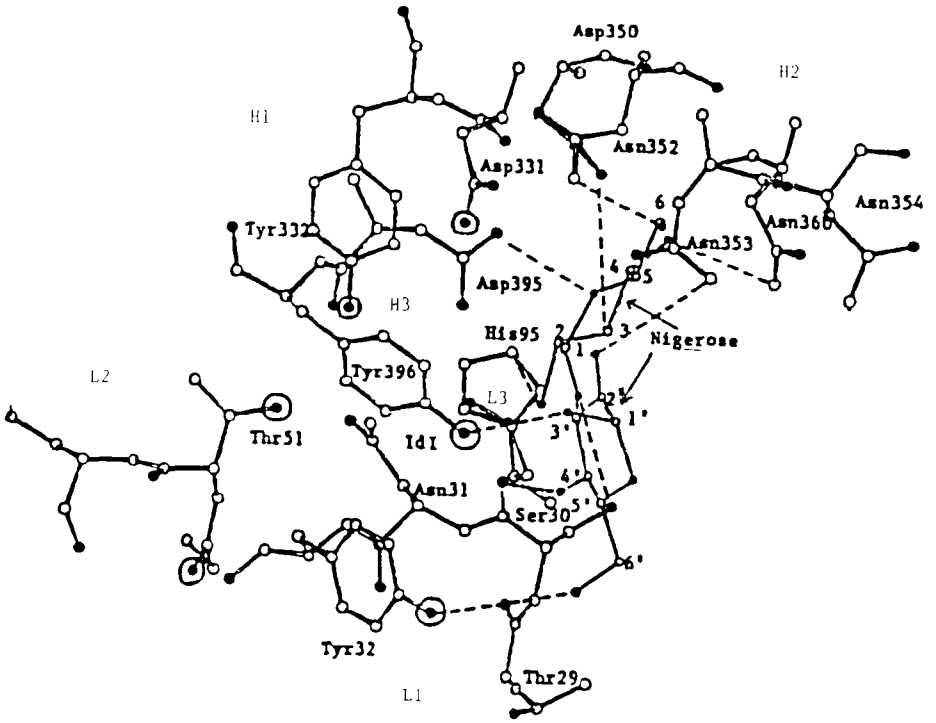

Fig. 2. The predicted antibody-combining site of immunoglobulin MOPC104E. CDRs are indicated as L1, L2, L3, H1, H2, and H3. A nigerose molecule can hydrogen bond with many of the CDR amino acid residues.

MOPC104E has not been crystallized, and our theoretically predicted atomic coordinates make up the only available 3-D structure.

If we look directly into the antibody-combining sites determined experimentally, CDRH3 and CDRL3 are located in the center, whereas CDRH2 and CDRL2 are on the periphery. Thus, in our predictive procedure, we started with CDRH3 and CDRL3 first and gradually worked our way from the center of the combining site toward the periphery. Indeed, the six CDRs were so tightly packed together that they formed a compact surface. The combining site of MOPC104E could easily fit precisely with a nigerose molecule (Fig. 2) with additional glucose units lying on the surface of the site (*34*). Similar predictive studies were undertaken by many other researchers (*see 35*).

Determining the 3-D structures of antibody-combining sites is of fundamental importance to our ultimate goal of designing therapeutic antibodies with required specificities. In principle, the resulting surface formed by a specified set of six CDRs will be able to interact strongly with a known antigen. However, our current knowledge of such surfaces on protein molecules is extremely limited. In fact, most of the docking studies between macromolecules are rather ill-defined. Some simple features are known. For example, if the antigen molecules has a protrusion, the antibody-combining site should have a depression to fit the protrusion; and

Table 1
Certain CDR Amino Acid Sequences Are Found in Antibodies with Different Specificities, Except CDRH3; thus, CDRH3 May Play a Unique Role in Conferring Fine Specificity, and the Other Five CDRs Can Contribute to Binding Affinity

CDR	Sequence	Number of different specificities
CDRL1	ARG SER SER GLN SER LEU VAL HIS SER --- ASN GLY ASN THR TYR LEU HIS	16
CDRL2	LYS VAL SER ASN ARG PHE SER	29
CDRL3	SER GLN SER THR HIS VAL PRO --- --- --- --- --- --- TRP THR	4
CDRH1	ASP TYR TYR MET ASN	10
CDRH2	ASP ILE ASN PRO --- --- ASN ASN GLY GLY THR SER TYR ASN GLN LYS PHE LYS GLY	8
CDRH3	TYR TYR TYR GLY SER SER --- --- --- --- --- --- --- --- --- --- --- --- LEU VAL	1

vice versa. Detailed studies of charge interactions, hydrophobic effects, and so forth, are definitely required in order to understand what a perfect fit between biological macromolecules is (*36*).

Because of the lack of this detailed knowledge, we have thus decided to collect all available amino acid and nucleotide sequences of antibodies with known specificities. From this collection, we are hoping that some CDR amino acid sequences may be correlated with the ability of binding certain antigen molecules. If we assume that all six CDRs contribute to specificity equally, then each CDR should be associated with many different specificities. For example, with 10^2 different sequences for each CDR, a cell can generate $(10^2)^6 = 10^{12}$ antibodies with different specificities, even without considering junctional diversities and other genetic mechanisms. Each CDR would be associated with $(10^2)^5 = 10^{10}$ different specificities, since the other five CDRs can vary randomly.

Our data base is not that large. Nevertheless, random assortment of CDRs seems to be the case for CDRL1, CDRL2, CDRL3, CDRH1, and CDRH2 with the limited number of antibodies with known specificities available for our study, as indicated in Table 1. For example, a certain amino acid sequence of CDRL1 can be found in many antibodies with different specificities. The same is also true for CDRL2, CDRL3, CDRH1, or CDRH2. In fact, sometimes, the same light-chain variable region amino acid sequence can constitute antibodies of completely different specificities. The only exception to this idea is CDRH3. A given amino acid of CDRH3 is nearly always associated with a unique antibody specificity (*37*).

In other words, the contributions of the six CDRs to antibody specificity are different. Currently, the idea is that CDRH3 may be responsible for fine specificities, whereas the other five CDRs can contribute significantly to binding affinities. To produce a good antibody requires matching CDRH3 perfectly with the antigen molecule. Contour fitting may be achieved by varying the length and amino acid residues of CDRH3, which is the only CDR with extensive length variations (*38*). Recently, a CDRH3 has been isolated from an anti-gp120 antibody, and found to have good binding specificities and affinities (*39*). Genetically, CDRH3 is coded by the D minigene or minigenes (*40*) together with the incorporation of N (*41*) and/or P segments (*42*). In most cases, V-D-J rearrangement of heavy chains precedes V-J rearrangement of the light chains, suggesting that a cell may have decided what antibody to produce at the pre-B-cell stage.

The collection of CDRH3 sequences associated with known antibody specificities is extremely informative (*43*). For humans, the length of CDRH3 varies from 2 to 26 amino acid residues with a mean of 11.6. The shortest ones were the result of direct V-J joining of heavy-chain genes (*44*). For mouse, the length variation is more restricted, from 2 to 19 residues with a mean of 8.7. For froglet and tadpole, the range is even smaller. Thus, human antibodies may be the most versatile ones. In principle, to design an antibody of a given specificity, we should be able to pick the possible CDRH3 sequences from this collection. These sequences may be directed toward different areas of the surface of the antigen molecule. For example, the three antilysozyme antibodies studies recently by X-ray diffraction studies have CDRH3 sequences of five, seven, and eight amino acid residues (*16,45,46*). The five-residue CDRH3 is directed toward a protruded portion of the lysozyme molecule, whereas the seven- and eight-resisue ones toward relative flat portions of the molecule. Other examples of CDRH3 length variations in the antibody-combining sites can be found in the Introduction of reference *15*, where a number of stereo-pairs of 3-D structures of antibodies are summarized.

Since our collection of such information and our knowledge of how compact biological macromolecular surfaces interact are very limited, the current practical approach of obtaining antibodies with desired specificities is to rely on the antibody-production talent of rodents. Antigens were injected into mouse or rat, and high-affinity antibody-producing cells were constructed into hybridoma cell lines. The genetic segments coding for the CDRs from these cell lines were then grafted onto selected FRs of human antibodies in order to minimize xenogeneic rejection (*47–49*). This techique of humanizing rodent antibodies for therapeutic use in human patients circumvents the understanding of how antigen and antibody molecules interact at their contacting surfaces.

This seemingly straightforward approach, however, encountered unexpected difficulties, possibly because of the fact that joining of rodent

CDRs with human FRs may distort the 3-D "sandwich" structure formed by the two β-sheets with antiparallel strands in the framework region. The humanized rodent antibodies can thus be unstable and have low binding affinities. The current belief is that the human FRs used for CDR grafting should be very similar to the original rodent antibody FRs in order to maintain activity and avidity. We have developed a search program to make such matches, which are essential for the success of making useful therapeutic humanized rodent antibodies.

Many years ago, we noted that some FR sequences are identical among human, mouse, and rabbit antibodies, and have probably been preserved for over 20 million years (*21*). During the process of FR sequence diversification among species, some of the ancient sequences are still functional. Indeed, these sequences can be useful in providing a tool for bridging the gap between human and rodent antibodies. In principle, a complete set of human light- or heavy-chain FRs can differ from that of mouse by only one amino acid residue (*50*). This shared framework region seems to be an ideal device in humanizing mouse antibodies for therapeutic uses.

During the past quarter century, our understanding of the structure and function of antibodies has advanced dramatically. Numerous scientists have contributed to this important field. We are just some of the fortunate ones being at the right place at the right time who have participated in finding a few pieces of the big puzzle. As the versatility of the antibody molecules becomes clearer, these protein molecules are called upon to do various tasks in this new world of genetic engineering. Their basis property of binding antigen molecules tightly provides a sensitive method of detecting small amounts of antigens. Therapeutic antibodies with reasonable half-lives in patients have been tried in various disease processes. Specific antibodies with catalytic activities similar to enzymes were discovered (*51*). All such studies can eventually contribute to our understanding of how biological macromolecular surfaces interact.

REFERENCES

1. Poulik, M. D. and Edelman, G. M. (1961), *Nature* **191,** 1-4.
2. Hilschmann, N. and Craig, L. C. (1965), *Proc. Natl. Acad. Sci. USA* **53,** 1403-1409.
3. Milstein, C. (1967), *Nature* **216,** 330-332.
4. Milstein, C. and Pink, J. R. L. (1970), *Prog. Biophys. Mol. Biol.* **21,** 209-263.
5. Franek, F. (1969), in *Symposium on Developmental Aspects of Antibody Formation and Structure,* Prague, pp. 311-313.
6. Kabat, E. A. (1967), *Proc. Natl. Acad. Sci. USA* **58,** 229-235.
7. Kabat, E. A. (1968), *Proc. Natl. Acad. Sci. USA* **59,** 613-619.
8. Wu, T. T. and Kabet, E. A. (1970), *J. Exp. Med.* **132,** 211-250.

9. Kabat, E. A., and Wu, T. T. (1971), *Ann. NY. Acad. Sci.* **190,** 382–393.
10. Kabat, E. A., Wu, T. T., and Bilofsky, H. (1976), *Variable Regions of Immunoglobulin Chains,* Bolt, Beranek and Newman, Cambridge, MA.
11. Wu, T. T. (1969), *Proc. Natl. Acad. Sci. USA* **63,** 400–405.
12. Gehring, K., Leroy, J.-L., and Gueron, M. (1993), *Nature* **363,** 561–565.
13. Poljak, R. J., Amzel, L. M., Avey, H. P., Chen, B. L., Phizackerley, R. P., and Saul, F. (1973), *Proc. Natl. Acad. Sci. USA* **70,** 3305–3310.
14. Schiffer, M., Girling, R. L., Ely, K. R., and Edmundson, A. B. (1973), *Biochemistry* **12,** 4620–4631.
15. Kabat, E. A., Wu, T. T., Perry, H., Gottesman, K., and Foeller, C. (1991), *Sequences of Proteins of Immunological Interest,* 5th ed., National Institutes of Health, Bethesda, MD.
16. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986), *Science* **233,** 747–753.
17. Baltimore, D. (1981), *Cell* **24,** 592–594.
18. Reynaud, C.-A., Anquez,V., Dahan, A., and Weill, J.-C. (1985), *Cell* **40,** 283–291.
19. Reynaud, C.-A., Anquez, V., Dahan, A., and Weill, J.-C. (1989), *Cell* **59,** 171–183.
20. Matsuda, F., Shin, E. K., Nagaeka, H., Matsumura, R., Haino, M., Fukita, Y., Taka-ishi, S., Imai, T., Riley, J. H., Anand, R., Soeda, E., and Honjo, T. (1993), *Nature Genetics* **3,** 88–94.
21. Kabat, E. A., Wu, T. T., and Bilofsky, H. (1978), *Proc. Natl. Acad. Sci. USA* **75,** 2429–2433.
22. Bernard, O., Hozumi, N., and Tonegawa, S. (1978), *Cell* **15,** 1133–1144.
23. Early, P., Huang, H., Davis, M., Calame, K., and Hood, L. (1980), *Cell* **19,** 981–992.
24. Siebenlist, U., Ravetch, J. V., Korsmeyer, S., Waldmann, T., and Leder, P. (1981), *Nature* **294,** 631–635.
25. Wu, T. T. and Kabat, E. A. (1982), *Proc. Natl. Acad. Sci. USA* **79,** 5031–5032.
26. Nishikawa, K. and Noguchi, T. (1991), *Methods in Enzymol.* **202,** 31–44.
27. Kabat, E. A. and Wu, T. T. (1972), *Proc. Natl. Acad. Sci. USA* **69,** 960–964.
28. Wu, T. T. and Kabat, E. A. (1972), *Proc. Natl. Acad. Sci. USA* **68,** 1501–1506.
29. Brookhaven Protein Data Bank, Brookhaven National Laboratory, Long Island, NY.
30. Honig, B., Kabat, E. A., Katz, L., Levinthal, C., and Wu, T. T. (1973), *J. Mol. Biol.* **80,** 277–295.
31. Wu, T. T. and Kabat, E. A. (1973), *J. Mol. Biol.* **75,** 13–31.
32. Stanford, J. M. and Wu, T. T. (1981), *J. Theor. Biol.* **88,** 421–439.
33. Coutre, S.E., Stanford, J. M., Hovis, J. G., Stevens, P. W., and Wu, T. T. (1981), *J. Theor. Biol.* **92,** 417–434.
34. Hovis, J. G. and Wu, T. T. (1985), in *New Methodologies in Studies of Protein Cofiguratioan,* Wu. T. T. (ed.), Van Nostrand Reinhold, New York, pp. 142–164.
35. Padlan, E. A. and Kabat, E. A. (1991), *Methods in Enzymol.* **203,** 3–20.
36. Bacon, D. J. and Moult, J. (1992), *J. Mol. Biol.* **225,** 849–858.
37. Kabat, E. A. and Wu, T. T. (1991), *J. Immunol.* **147,** 1709–1719.
38. Chothia, C. and Lesk, A. M. (1987), *J. Mol. Biol.* **196,** 901–921.

39. Levi,M., Sallberg, M., Ruden, U., Herlyn, D., Maruyawa, H., Wigzell, H., Marks, J., and Wahreu, B. (1993), *Proc. Natl. Acad. Sci. USA* **90,** 4374–4378.
40. Kurosawa, Y. and Tonegawa, S. (1982), *J. Exp. Med.* **155,** 201–218.
41. Alt, F. and Baltimore, D. (1982), *Proc. Natl. Acad. Sci. USA* **79,** 4118–4122.
42. Lafaille, J. J., DeCloux, A., Bonneville, M., Takagaki, Y., and Tonegawa, S. (1989), *Cell* **59,** 859–866.
43. Wu, T. T., Johnson, G., and Wu, T. T. (1993), *Proteins: Structure Genetics and Function* **16,** 1–7.
44. Lewis, S. M., Hesse, J. E., Mizuuchi, K., and Gellert, M. (1988), *Cell* **55,** 1099–1107.
45. Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., and Davies, D. R. (1987), *Proc. Natl. Acad. Sci. USA* **84,** 8075–8079.
46. Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., and Davies, D. R. (1989), *Proc. Natl. Acad. Sci. USA* **86,** 5938–5942.
47. Jones, P.T., Dear, P. H., Foote, J., Neuberger, M. S., and Winter, G. (1986), *Nature* **321,** 522–525.
48. Verhoeyen, M., Milstein, C., and Winter, G. (1988), *Science* **239,** 1534–1536.
49. Co, M. S., Deschamps, M., Whitley, R. J., and Queen, C. (1991), *Proc. Natl. Acad. Sci. USA* **88,** 2869–2873.
50. Wu, T. T. and Kabat, E. A. (1992), *Mol. Immunol.* **29,** 1141–1146.
51. Shokat, K. M. and Schultz, P. G. (1991), *Methods in Enzymol.* **203,** 327–351.

DISCUSSION

T. T. Wu

Paul: Does direct evidence exist linking the length of CDRH3 or other CDRs with binding affinity?

Wu: The only thing that I have any knowledge of is binding constants for lysozymes and antilysozyme Fab fragments. The original one, the D 1.3, binds lysozyme with, I think, a binding constant of $5 \times 10^7 M^{-1}$ (*1*). Subsequently, there are two additional ones from David Davies' group with binding constants on the order of $10^{10} M^{-1}$ (*2,3*). In that particular case, the antibody actually has a very short CDRH3, but the binding constant turns out to be quite large. So, I do not think you can link the binding constant to one particular CDR. The binding constant probably is related to all six CDRs.

Zouali: We have made a computer search of the length of the CDRH3s of human antibodies and, we have compared polyreactive human antibodies to those that are monospecific. We found that there is no significant difference in the length of the CDR3 of the heavy chain.

Polanovsky: What is your opinion about humanization of mouse antibodies? Is this a promising and general strategy?

Deyev: It is now possible to isolate directly human antibody V genes and to prepare single-chain antibodies of human origin.

Wu: If you wanted to have a gigantic library of all different kinds of human sequences and then express it in various kinds of expression vectors, like the phages or the fungus, then you would have to go through a big screening procedure. On the other hand, if consistent human hybridoma techniques can be developed and if these hybridomas grow out very nicely, that is another way to go. I think the reason that the humanizing of mouse antibodies worked out very nicely is that because you can actually pick out a specific antibody for a specific antigen relatively quickly. Of course, other techniques are being developed simultaneously, like the phage expression of massive amounts of random sequences.

DISCUSSION REFERENCES

1. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986), *Science* **233,** 747–753.
2. Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., and Davies, D. R. (1987), *Proc. Natl. Acad. Sci. USA* **84,** 8075–8079.
3. Padlan, E. A., Silverton, E. W., Sheriff, A., Cohen, G. J., Smith-Gill, S. J., and Davies, D. R., (1989), *Proc. Natl. Acad. Sci. USA* **86,** 5938–5942.