Antibodies (II): Another Look at the Diversity Problem

The origin of antibody diversity—that is, how to account for the ability of a single animal to make as many as a million different antibodies—is a scientific puzzle of long-standing. The answer to this puzzle undoubtedly resides in the genes coding for the protein chains of which antibody molecules are composed. But for many years, researchers who were looking for the answer had to content themselves with examining the structures of a large number of the proteins and then reasoning backward to work out the gene arrangements that could best account for the structures.

Now, however, as a result of the development of new techniques for probing genes and determining their nucleotide sequences, investigators are acquiring the kind of information about the genes that they need to unravel the mysteries of antibody diversity. At the same time, the protein chemists, aided by more efficient methods for determining and comparing the amino acid sequences of proteins, have expanded the repertoire of known antibody chain structures.

Both the gene and protein work suggest that there may be more antibody genes than was thought a few years ago. In addition, the gene structures have suggested some plausible—if not proven—mechanisms for generating still more diversity from an already large population of antibody genes. All in all, there is a growing consensus that antibody diversity may be generated by a combination of several processes, although individual investigators differ widely in their views of which of these processes is primary and which are secondary.

The polypeptide chains of which antibody molecules are composed consist of a variable region, which differs from one chain to the next, and a constant region, which is the same for all chains of the same type (Fig. 1). The genetic information for the constant regions is encoded separately from that for the variable regions. Investigators now know there is only one copy of each kind of constant region gene. Thus, the diversity issue only concerns the generation of the different variable regions.

Three major competing theories currently seek to explain this phenomenon. According to the germ-line theory of antibody diversity, the variability is encoded by large numbers of variable region genes present in immature (germ) cells and inherited by all antibody-producing cells. The somatic mutation theory holds that relatively few variable region genes in the germ cells are inherited by antibody-producing cells and that diversity is generated by mutations in these genes. A third kind of model postulates that rearrangement of variable gene segments during the differentiation of the antibody-producing cells generates antibody diversity.

The gene rearrangement models were proposed to account for certain observations on the nature of variable region structures. Several years ago, Elvin Kabat of Columbia University Medical School and a consultant to the National Cancer Institute (NCI) and Tai Te Wu of Northwestern University noted that amino acid differences were more likely to occur in some portions of the variable regions of antibody chains than in others. These regions are called hypervariable and the less frequently altered segments are called framework regions. Kabat and Wu predicted that the amino acids of the hypervariable segments form the antigen-combining sites of antibodies, a prediction that was subsequently confirmed when x-ray crystallographers determined the exact three-dimensional structure of antibodies.

In addition, Kabat and Wu proposed that the DNA segments coding for the hypervariable regions are separate from those coding for the framework amino acids. The idea was that the hypervariable segments could be inserted individually between the framework regions to form a complete variable region gene in antibody-producing cells.

Meanwhile, Donald Capra of the University of Texas Health Science Center at Dallas and Thomas Kindt of the National Institute of Allergy and Infectious Diseases had come to the conclusion that neither the germ-line theory nor the somatic mutation theory were adequate to explain everything known about antibody structures and how they are inherited. Capra and Kindt also proposed that the hypervariable region sequences are encoded separately from the framework structures and then joined in various combinations to produce active variable region genes.

It is safe to say that few immunologists were enthused about the gene rearrangement models. Kabat himself says that no one believed their original model because there were no known mechanisms in mammalian cells by which such insertions could be achieved. Moreover, when Susumu Tonegawa of the Basel Institute for Immunology and Walter Gilbert and Allan Maxam of Harvard University determined the nucleotide sequence of a variable region gene from mouse embryo cells (*Science*, 20 October, p. 298), they found the hypervariable sequences already in place, a result not compatible with Kabat and Wu's suggestion that those sequences were not present in germ cells but are inserted between the framework regions only in mature cells. Capra, incidentally, points out that he never said that the hypervariable regions were missing in the embryonic genes; he thinks that they might be present there, although exchangeable for alternative segments.

In response to the recent findings on antibody gene structure and the new developments regarding the existence of intervening sequences in genes, Kabat and Wu with Howard Bilofsky of Bolt Beranek & Newman Inc., have proposed a modification of their earlier theory. They now suggest that in embryonic genes intervening sequences may be found between the nucleotide segments coding for the framework regions. The complete variable region gene would then be assembled by eliminating the intervening sequences and inserting hypervariable "minigenes" between the framework segments. The Basel and Harvard researchers have found no indication of such intervening sequences in variable region genes from mouse embryo cells, however.

But some investigators have questioned whether cells from 12-day-old mouse embryos, like those used by Tonegawa, are truly germ-line cells. They think some differentiation may have occurred by this stage of development. Kabat suggests that to verify or disprove the rearrangement theory, it might be necessary to examine antibody gene arrangements in DNA from egg and sperm cells, which certainly should be germ-cell DNA. But Tonegawa has now compared DNA from mouse sperm with that from embryos by gene-mapping techniques and has found no differences in the patterns obtained for the two kinds of DNA, a result that appears to disprove Kabat, Wu, and Bilofsky's suggestion that all

0036-8075/78/1027-0412$00.50/0

antibody diversity can be generated by gene rearrangement.

Nevertheless, there are strong indications that gene rearrangements may make some contribution to antibody diversity. The gene studies of Tonegawa and his colleagues have shown that the variable regions of mouse light chains are encoded by two separate gene segments. One of the segments encodes amino acids 1 to 97 or 98, and the other encodes the remaining 13 or so amino acids of the variable region. (This short segment is now called the J region because it joins the variable and constant regions of antibody chains.)

Several investigators, including Martin Weigert of the Institute for Cancer Research, Lee Hood of the California Institute of Technology, and Michael Potter and Stuart Rudikoff of NCI, who have been determining the amino acid sequences of light chains, have also concluded that the variable regions are encoded by two individual gene segments that may combine randomly. Weigert and Hood have identified some 11 different J region segments, each presumably encoded by a different gene segment, in only one group of κ chains.

The third hypervariable region of mouse κ chains may include amino acids 100 and 101. If these are proved by gene sequencing to be encoded in the J region gene, as is now thought, rearrangement of at least the variable and J region genes may contribute to antibody diversity. Moreover, the variable regions of mouse heavy chains also appear to be composed of two distinct segments, according to Rudikoff and his colleagues.

At this time, however, many investigators do not think that all—or even most—of the variability can be accounted for by such rearrangements. These researchers are more interested in determining whether germ-line gene number or somatic mutations make the greater contribution.

A few years ago, the evidence seemed to favor the view that there were very few germ-line antibody genes and, consequently, that somatic mutations explained almost all of antibody diversity (*Science*, 1 November 1974, p. 432). But more recent data have suggested to some investigators that the reports of the death of the germ-line theory have been premature, at least in regard to κ chains, which account for around 95 percent of all light chains in mouse antibodies. The remainder are λ chains for which the consensus in favor of somatic mutation still holds.

The evidence concerning the number of κ chain variable genes comes from

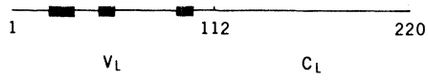

Fig. 1. Diagram of the structure of an antibody light chain consisting of about 220 amino acid residues. The variable region (V_L) is generally thought to extend from amino acid 1 to amino acid 112 and the rest of the chain is the constant region (C_L). The heavy bars indicate the approximate positions of the hypervariable regions where the frequency of amino acid differences between chains is very high.

both gene-counting studies and analysis of the amino acid sequences of light chains. One way to count the genes directly is to make a radioactive DNA probe that is complementary to the nucleotide sequences of the genes and will thus stick to any DNA fragment bearing them. The probe is usually a DNA copy of the messenger RNA (mRNA) corresponding to a particular light chain. The mRNA can be obtained from cells that produce the chain in question.

Philip Leder, Jonathan Seidman, and their colleagues at the National Institute of Child Health and Human Development (NICHD), have used this approach to count the genes corresponding to two different κ chains. Based on differences in their amino acid sequences, the two light chains belong to different κ chain groups. By Leder's definition two κ chains belong in the same group if at least 90 percent of their amino acid residues are identical and if the differences that do exist between them are mainly in the hypervariable regions. The reason for grouping light chains according to the relatedness of their amino acid sequences is to get a rough idea of the number of genes needed to code for the proteins. If two or more proteins are very similar and thus members of the same group, then they might be coded for by one gene with mutations accounting for the few differences in their sequences.

But the NICHD workers are finding that several genes exist for each group. The probe used to search for variable region genes belonging to one of the κ chain groups picked out six DNA fragments bearing variable region genes in a DNA preparation from a line of antibody-producing cells; the other probe stuck to seven or eight such fragments.

According to Leder and Seidman, embryonic cells also contain similar patterns of about six genes per group, a result suggesting the genes are present in germ-line cells as well as in mature ones.

The close structural relation between the genes within one group was confirmed by determination of some of their nucleotide sequences. Leder and his colleagues first cloned two of the cellular DNA fragments detected by one of the probes. The fragments each carried a variable region gene, and the nucleotide sequences of the two genes turned out to be very similar, but not identical. These two genes are not actually expressed in the cell line from which they were obtained, but their sequences closely resemble that of the gene that is expressed. This was expected because the probe by which they were identified is a copy of the mRNA produced by the cell line and thus is the equivalent of the expressed gene.

Additional results suggesting that there may be six or so genes for every κ chain group come from studies of the amino acid sequences of the proteins. Studies of the $V_\kappa 21$ group of light chains, which have been carried out by Weigert and Hood and also by Potter, have been especially revealing. At present, 18 of the 22 $V_\kappa 21$ proteins with known amino acid sequences can be divided into at least six subgroups, each of which appears to require a distinct variable region gene. The researchers think that each subgroup has its own gene because if this were not the case, several independent but identical mutations occurring in a germline gene in different cell lines would be needed to generate the two or more members of a given subgroup. This is more coincidence than geneticists normally like to invoke. Moreover, Weigert, in collaboration with Oscar Valbuena and Robert Perry, also of the Institute for Cancer Research, in some indirect gene-counting experiments have also found evidence for the existence of six genes coding for $V_\kappa 21$ light chains. The investigators thus conclude that there is one gene for each of the six subgroups and that the different members of a subgroup are generated by somatic mutation of the appropriate gene.

But four of the $V_\kappa 21$ proteins did not fit into any of the six subgroups. Designation of a new subgroup requires that it have a minimum of two members, and since partners have not been found for any of the four unclassified chains, they cannot now be said to represent additional $V_\kappa 21$ subgroups. If partners turn up as more proteins are sequenced, however, then the existence of additional subgroups may be established. In that case, investigators are going to have to figure out the reason for the discrepancy between the numbers of $V_\kappa 21$ genes counted by direct and indirect methods.

Athough there is a growing agreement about the existence of a half-dozen or so genes for each group of κ chains, important questions remain about whether

there are enough of the genes to explain most or all of antibody diversity or whether somatic mutation, which is known to occur, is the primary mechanism underlying the diversity. Whole antibody molecules are formed of two identical light plus two identical heavy chains. If the combination is random, then 1000 different genes for light chain variable regions plus another 1000 for the heavy chains would be sufficient to produce 1 million different antibody molecules, which is usually cited as the minimum number that an individual animal is capable of producing.

Speaking of Science

Soviet Science: A Wonder Water from Kazakhstan

Soviet scientists seem to have an unusual fascination with abnormal states of water. Only 5 years ago, Soviet physicist Boris V. Derjaguin resolved a controversy he had helped to create by demonstrating that "polywater"—an anomalous form of water then considered to be polymeric—was simply water containing a large concentration of dissolved minerals. Now, according to the Information Office of the Soviet embassy, twin brothers at a relatively obscure research institute in Kazakhstan have identified another form of water that appears to have a greater biological activity than ordinary tap water.

The brothers, Vadim and Igor Zelepukhin of the Institute of Fruit-Growing and Vine-Growing of Kazakhstan, made their discovery while investigating the biological properties of freshly melted snow. Soviet scientists have known for some time that fresh meltwater has the capacity to stimulate some biological processes. It has been theorized that the meltwater retains some of the order that is characteristic of frozen water and that this increased order alters vital reaction rates within cells; some American investigators have theorized, similarly, that water within cells is more highly ordered than ordinary water and that this increased order is essential to proper control of enzymes.

The brothers Zelepukhin were studying this phenomenon by observing the uptake of water by cut leaves. They would weigh leaves, place them in tap water, boiled water, or fresh meltwater for 1 hour, and then reweigh them to determine their water uptake. Typically, they found that the leaves absorb two to three times as much meltwater as either tap water or boiled water. In one series of experiments, however, Vadim inadvertently overturned several boxes containing boiled water. To save the experiment, he quickly boiled more water and cooled it under the tap to the required temperature of 20°C. Suprisingly, the Zelepukhins found that leaves absorbed five to six times as much of this water as of ordinary tap water and two to three times as much as meltwater.

After confirming this result in other types of leaves, they tried other experiments, such as using the "bioactive" water in place of tap water to soak seeds before they were planted. In one study, for example, they found that cotton plants grown from such seeds yielded 10 to 12 percent more cotton than the farm's average. In all stages of the plants' development, Vadim Zelepukhin says, "the experimental plants were superior to the control plants in all physiological characteristics. There was a more intensive chlorophyll formation in the leaves, a more active water exchange, and so forth." The quality of the fiber was also better.

Greater yields were also obtained when they soaked tomato, potato, maize, and wheat seeds in the water. Particularly responsive, Zelepukhin says, were sugar beets. Root weight increased 40 percent and sugar content 1.5 percent. They also observed that bioactive water stimulates root formation in fruit tree cuttings better than either tap water or water containing the root stimulant heteroauxin.

After many experiments, the brothers finally concluded that the activity of the water results from the fact that it has been thoroughly degassed. Boiling removes dissolved gases and the quick cooling prevents them from redissolving immediately. They were thus able to reproduce their results by degassing water with a saturation syringe. They were also able to show that water prepared either with the syringe or by boiling and quick cooling loses its stimulatory power if it is exposed to the atmosphere for more than an hour or two, but retains the activity if it is stored in a sealed container. The biological activity of meltwater can be explained in much the same fashion: Freezing degasses water, although to a lesser extent than boiling. The biological activity of meltwater is thus not as great as that of water boiled and quickly cooled.

Degassed water differs considerably from tap water in its physicochemical properties, says Igor Zelepukhin. Its conductivity is decreased considerably and there are increases in its density, viscosity, surface tension, energy of intermolecular interaction, and internal pressure. Degassed water thus bears a closer resemblance to the fluid in cells than does tap water, he adds, and this could account for its effects.

In some other experiments conducted at the Kazakhstan branch of the Research Institute of Dietotherapy of the Academy of Medical Sciences of the U.S.S.R., the Zelepukhins gave degassed water to white rats for a month and tap water to a control group. The most striking difference between the groups, they say, was an increase in the average hemoglobin level from 75.8 units in controls to 93.4 units in the experimental animals. The test animals also exhibited enhanced activity for most enzymes and had only insignificant quantities of lactic acid in their muscles—an indication, Zelepukhin says, of their greater viability and decreased fatigability under equal working conditions. They are currently experimenting with farm livestock and expect to achieve good results since other investigators have already shown that cattle, for example, gain weight faster when given fresh meltwater.

In one other application of note, the construction industry of the Soviet Union has begun to use degassed water on an experimental basis. They have observed that concrete prepared with it is 8 to 10 percent stronger than concrete prepared with ordinary water. The water is obtained relatively easily by use of a mobile milk-pasteurizing unit, in which milk is heated and then quickly cooled in a special radiator. Soviet investigators predict that a variety of other uses will soon be found for this "magical" material.

—THOMAS H. MAUGH II

0036-8075/78/1027-0414$00.50/0

Estimates vary for the total number of variable region genes for κ chains. One of the higher estimates is that of Weigert and his colleagues. Based on a statistical analysis of the known amino acid sequences of κ chains, they concluded that the proteins could be subdivided into about 50 groups. If six or seven genes are needed to code for the members of each group, there should be about 350 variable region genes for the κ chains. This is more than previously thought but still far short of 1000. Weigert says he now thinks that the number of germ-line genes is sufficient to account for only about 10 percent of antibody diversity, with other mechanisms providing for the rest. These other mechanisms would include somatic mutation and combination between a number of J and variable region genes.

Some investigators, including Potter, do not think that all the κ chain groups have been identified, however. He has identified a number of new lines of antibody-producing cells from an inbred strain of mice and says that it is easy to find new light chains. If the number of groups continues to grow, then the magic number of 1000 may yet be reached. Potter agrees that somatic mutation generates some antibody diversity but relegates it to a secondary role compared to that of germ-line gene number.

Although Leder has come to a similar conclusion, he has recently suggested a novel mechanism for generating additional diversity. He bases the hypothesis on the results of the gene-sequence studies being carried out in his laboratory. In addition to finding a high degree of similarity in the nucleotide sequences of the two variable genes they have analyzed, the NICHD group has determined that the similarities extend well beyond the genes themselves into the nucleotides on either side of the genes.

This close structural resemblance over a long stretch of nucleotides would facilitate association between related genes carried on the two members of a pair of chromosomes. These associations may occur during cell division and when they do, the two genes may recombine, that is, exchange nucleotides with one another. Such exchanges by themselves can generate alterations in gene structure and thus increase diversity. Moreover, during the exchange, mistakes may happen with the potential for producing still more variability.

Leder points out that the situation regarding the antibody genes presents a striking contrast to that observed with the two genes coding for one of the proteins that make up hemoglobin. Mutations in the genes for this protein, which may alter or even destroy its function, are poorly tolerated. According to Leder and his colleagues, the nucleotide sequences flanking these structural genes are very dissimilar, a situation that would tend to minimize recombination between them and consequently reduce the likelihood of the genes being altered.

If recombination is one of the mechanisms for generating antibody diversity, then a way to prevent catastrophic gene loss, which might result from extensive recombination between dissimilar groups of genes, would be needed. Loss of variable genes could deprive the animal of needed antibodies. Leder suggests that such losses could be prevented if the ability to recombine were restricted to the few variable region genes within a single group. And this may be the case. Experiments in his laboratory indicate that similarities in the flanking sequences are restricted to members of the same group. This would mean that recombination would occur primarily among the six or so genes belonging to one group and not between genes of different groups. Such an arrangement would make catastrophic loss of many variable region genes unlikely.

Although Tonegawa has also observed extensive similarities in the nucleotide segments flanking variable genes of κ chains, he is still cautious about the significance of the findings. In this context, he thinks a mechanism like that proposed by Leder is possible but as yet unproved.

Although all the issues concerning antibody diversity have not yet been resolved, current research is illuminating the mechanisms by which the immune system can generate the multitude of antibodies it needs to deal with the essentially unlimited number of antigens it encounters. It now appears that multiple germ-line genes, somatic mutations, and gene rearrangements are all involved to some degree. No one expected that a system this complicated would be easy to unravel. And it has not been. But at last the antibody diversity problem is yielding to techniques that are becoming increasingly capable of tackling a question of this magnitude.—JEAN L. MARX

UPDATE

Grapes Inactivate Viruses, but Not in Body

Wine lovers rejoiced last year when Jack Konowalchuk and Joan I. Speirs of Health and Welfare Canada demonstrated that grape juice and wine can inactivate viruses in test tubes (*Science*, 3 June 1977, p. 1074). New results from Dean O. Cliver and Kenneth D. Kostenbader, Jr., of the Food Research Institute at the University of Wisconsin at Madison, however, suggest that the antiviral effects of grape products may be healthful only if used to bathe wounds to prevent infections.

In studies supported by the Concord Grape Association and Welch Foods, Cliver and Kostenbader replicated the experiments of the Canadian investigators and found that grape juice does, indeed, inactivate several types of viruses in the test tube. But they also found that the inactivated viruses were reactivated when exposed to biological systems.

The Wisconsin investigators have been studying viral infections of the gastrointestinal tract in baby pigs maintained on food normally given to human beings; this system closely mimics the human gastrointestinal system. After killing a small number of the piglets, the investigators mixed the contents from six different sections of the digestive tracts with viruses that had been inactivated by grape juice. After 30 to 60 minutes of exposure, they found that 50 to 75 percent of the virus particles had been reactivated. They also mixed inactivated viruses with human blood serum and observed that 94 percent of the virus particles were reactivated. They thus conclude that ingested grape juice probably provides no protection against viruses.

The studies have not been a total loss, however. It is clear, Cliver says, that some relatively small molecule from grapes—Konowalchuk thinks it is phenol or its derivatives—is binding to the surface of the virus. This chemical prevents as many as 80 percent of the virus particles from binding to normal cells, and renders them almost completely noninfective. It could thus be a valuable tool, Cliver says, for investigating the mechanism by which viruses infect human cells, and may provide a lead to other agents that could produce irreversible inactivation. In the meantime, though, neither the Concord Grape Association nor Cliver recommends bathing wounds in grape juice or wine. Other topical disinfectants are more effective and have the added advantage of not staining the user purple.—T.H.M.

0036-8075/78/1027-0415$00.50/0