Antibodies Without Immunization

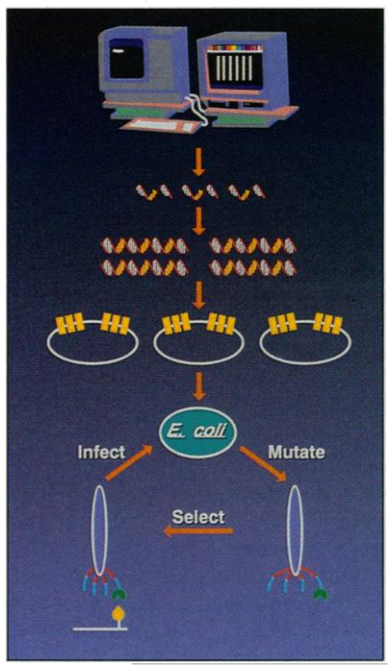
Richard A. Lerner, Angray S. Kang, Jim D. Bain, Dennis R. Burton, Carlos F. Barbas, III

Immunization of animals to produce antibodies for laboratory and medical use is among the time-honored processes in science. However, recent experimental data have indicated that this process might be replaced with a synthetic one. Examination of how naturally occurring antibodies are formed reveals some critical features and mechanisms that should be considered during the design of a replacement process: (i) the immune system is combinatorial, (ii) the processes of replication and recognition are linked, (iii) the antibody gene pool is largely naïve, and (iv) the system is capable of somatic refinement.

Copying and even improving this natural complexity requires the replication of the above features in a set of linked transactions. The ultimate goal is to provide a library of antibodies with nearly endless binding specificities and without dependency on animals or their cells.

The natural antibody system is fundamentally combinatorial. Diversity is generated by combinations of genes encoding two presumably randomly associated protein chains (heavy and light), each of which contains three hypervariable or complementarily determining regions (CDR's). The vast diversity of the antibody repertoire is achieved by the combinatorial association of these six CDR's in three-dimensional space. In a synthetic system the main hurdle to replacing the process of immunization is the need to mimic in vitro this large combinatorial process. This task is the opposite to that presented by cloning because diversity rather than homogeneity must be achieved. A further complication in antibody synthesis is that the CDR's are present on two different protein chains.

Earlier work on the assembly of immunoglobulins in *Escherichia coli* (*1–3*) and on amplification of the individual genes by the polymerase chain reaction (PCR) (*4–6*) provided a basis for the in vitro generation of antibody diversity. However, progress in the field was hampered by the assumption on the part of investigators that the probability of obtaining useful antibodies by random chain association was too low. Nevertheless, it has now been shown that the combinatorial problem could be solved by random association and expression of large numbers of genes encoding heavy and light chains in phage systems (*7*). This method is efficient for immunized animals because, as exposure to antigen continues, the induced repertoire becomes more homogeneous and results in an increased concentration of chains that can be paired in a way appropriate for antigen binding (*8*). Further, it may not be necessary to recapitulate the chain pairing found in the immunized animal. In fact, diversity can often be increased by a process of chain shuffling (*9*). A solution to the combinatorial problem cleared the main obstacle to replacing immunization in that an antibody library with a potential diversity exceeding that of intact animals could be contained in a volume of only a few microliters. It then remained necessary to obtain the starting gene pool, access large libraries, and improve their binding specificities.



Synthetic generation of antibodies. Sequences for antibody chains are constructed from synthetic nucleic acids randomized in the antigen binding region. These sequences are cloned into bacteriophages, which are then subject to mutation and selection.

Biological systems are especially powerful because recognition and replication are linked. Linkage can be at the level of whole organisms, individual cells, or viral entities. In the immune system the linkage exists because cells that recognize antigens at their surfaces are stimulated to divide. Screening of combinatorial antibody libraries is aimed at approximating this central feature of biological systems.

Systems in which diverse peptide sequences are displayed on the surface of filamentous bacteriophages (*10, 11*) have proved attractive for linking combinatorial antibodies with the genes that encode them in a package capable of recognizing antigen and replicating. This has been done in two ways (*12–14*). In each case both antibody chains (heavy and light) are linked to gene fragments that encode signals that direct them to the periplasmic space of *E. coli*. The genes for the antibody heavy chains are fused with either the phage gene III or gene VIII, which encode low and high copy phage surface proteins, respectively. This design allows heavy chains linked to phage proteins to encounter free light chains in the periplasmic space during the assembly of phage particles. Since the volume of an *E. coli* bacterium is somewhat less than 10^{-15} liter, a few hundred molecules of the antibody chains in the periplasmic space exceed the association constant of the heterodimer, thereby leading to appropriate assembly of the two-chain antibody structure on the phage surface. The difference between the methods is that in one the antibody genes are incorporated into the phage genome whereas in the other method they are in a smaller gene (phagemid), which is copied during viral replication. We use the phagemid (*13, 14*) because it readily permits gene cloning and its small size allows the generation of very large libraries. Furthermore, the phagemid–gene III constructions have yielded monovalent phage, thereby facilitating the selection of the highest affinity antibody clones.

Phage with specific antigen-binding properties can be selected and enriched by repeated attachment to, and elution from, a source of immobilized antigen. Although a concentration of 10^{11} phage per 50 microliters is possible, the transforming efficiency of the DNA limits the process so that in practice only 10^{8} different antibodies can be readily studied. Nevertheless, with these methods of screening combinatorial libraries, many antibodies to haptens and proteins from immunized animals and man have already been prepared.

With combinatorial libraries available, attention turned to circumventing the process of immunization. A primary consideration was the source of the antibody genes that make up the library (*15*). If educated

R. A. Lerner, A. S. Kang, J. D. Bain, and C. F. Barbas, III, are in the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037. D. R. Burton is in the Department of Immunology, The Scripps Research Institute, La Jolla, CA 92037.

libraries are used, the investigator is engaged in the formidable process of protein remodeling. Imagine the attendant difficulty of selecting a chain combination that must bind a hydrophobic hapten deeply within a binding pocket from a pool of antibody chains largely evoked to maximize the area of interacting protein surfaces to bind a protein antigen.

Naïve combinatorial libraries require a source of genes. Immunoglobulin genes from animals do not constitute a naïve gene pool because the animals have received many antigenic challenges. Furthermore, an animal repertoire may be evolutionarily restricted. For example, evidence is emerging that in humans the use of antibody chains is far from random. A further complication is the likely need for PCR to amplify sequences from natural sources because diversity decreases on subsequent rounds of gene amplification, a situation that may be favorable in immunized but not unimmunized animals.

Synthesizing combinatorial libraries then became the method of choice. Here some or all of the CDR's are encoded by synthetic nucleic acids randomized at most positions (*16*). Such libraries have a potential diversity greater than that of animals. We have selected high-affinity antibodies to fluorescein from combinatorial libraries in which the CDR3 regions of both the light and heavy chains are encoded by synthetic gene segments (*16*). Analysis of codon usage showed that selection for binding was attributable to the protein and was not due to some bias either in nucleic acid synthesis or in clonal expansion during phage selection (*16*). The selected antibodies showed consensus sequences related to both the general structural requirements of the antibody molecule and features required for specific binding to fluorescein. For example, only antibodies that had aspartic acid in position 101 of the heavy chain CDR3 were selected (*16*). In naturally occurring antibodies, the carboxylate of Asp^{101} often has an important structural role in that it can form a salt bridge with the guanidinium moiety of a framework arginine residue, thereby constraining the conformation of the CDR's (*17*). Thus, the in vitro selection system is mimicking a feature that results from natural selection and showed that antibody chains with synthetic CDR's can form the basis of large combinatorial libraries. Although selection for general structural features of the antibody molecule occurred when mixtures of synthetic and natural CDR's were combined, even this requirement may be relaxed when all the CDR's are synthetic. Thus, the system may have enough degrees of freedom to generate useful CDR combinations in a way not possible unless all the combinatorial parameters can vary in concert.

In some instances it will be necessary to mimic the natural somatic refinement process to increase the affinity of antibodies selected in the primary screening. We have been able to select for mutated antibodies of higher affinity by iterative procedures in which mutations are introduced by error-prone PCR (*18*). Although this staggered process of mutation and selection differs from the natural integrated selection scheme for cells with the highest binding affinity, the number of independent binding events that are available for study may be higher than those of natural systems. Whether sorting through large numbers of binding events can approximate or even exceed affinities obtained by more tightly integrated mutation and selection remains to be seen.

Apparently there is no longer any major technical obstacle to eliminating the process of immunization or even the use of animals for the preparation of antibodies. Whether or not the technology is general and applicable to all classes of molecules, we do know that principles learned for haptens have been generalizable to protein antigens. We anticipate that soon master combinatorial antibody libraries will be available for distribution and that such libraries will depend on the use of synthetic CDR's in which all positions are randomized except those known to be conserved in natural antibodies. Also, the length of the CDR's will be varied and will include some that are longer than those of natural antibodies. Only for human antibodies intended for therapy need considerations of potential immunogenicity temper any search for novel structures.

Combinatorial antibody libraries should be useful for isolating large enough numbers of antibodies to search for consensus sequence motifs implicated in binding to a particular class of antigens. This approach appears promising for generating antibodies that neutralize HIV. Here, highly specialized libraries may be constructed where motifs known to affect virus neutralization are incorporated into the synthetic CDR's. The master library should grow by acquiring these specialized libraries. Hence, the scientific community should adopt a technology that is sufficiently uniform for new constructs to be easily incorporated into the master library.

The use of master libraries could be compared to the process of immunization. The initial selection is like the primary response in which the affinity is expected to be variable and highly dependent on the chemical nature of the antigen. If the affinity constant is not suitable, mutagenesis and reselection, a process related to somatic refinement during sustained immunization, could be instituted. The iterative selection procedures used for screening and even mutating combinatorial libraries lend themselves to automation, and we can expect to see machines that carry out many of the steps.

Combinatorial antibody libraries were generated initially to match the chemical component of antibody catalysis with a protein component that was more understandable in chemical terms. Synthetic combinatorial antibody libraries have a diversity that is not subject to in vitro editing processes such as tolerance. This last feature allows selection for rare antibody species like those that may be required for acid-base catalysis or those that operate on structures of low immunogenicity, such as nucleic acids. In addition, unlike induction in animals that selects for binding events under aqueous conditions and neutral pH, antibodies can be selected from synthetic libraries that bind under various reaction conditions, including in the presence of organic solvents. Even antibodies that make covalent interactions can be selected by elution with nucleophiles. When known protein binding motifs are used, antibody chains that incorporate cofactors and metals can be included in the combinatorial library and thus extend the range of antibody catalysis to chemical reactions beyond the capability of amino acid side chains. Finally, master libraries can be thought of as a uniform and reproducible source of antibodies, a feature that can increase the precision of some immunological techniques and at the same time be of service to chemistry, biology, and medicine.

REFERENCES AND NOTES

1. M. A. Boss, J. H. Kenten, C. R. Wood, J. S. Emtage, *Nucleic Acids Res.* **12**, 3791 (1984).
2. S. Cabilly *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3273 (1984).
3. M. Better, C. P. Chang, R. R. Robinson, A. H. Horwitz, *Science* **240**, 1041 (1988).
4. J. W. Larrick, Y. L. Chiang, R. Sheng-Dong, G. Senyk, P. Casali, *Prog. Biotechnol.* **5**, 231 (1988).
5. R. Orlandi, D. H. Gussow, P. T. Jones, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3833 (1989).
6. L. Sastry *et al.*, *ibid.*, p. 5728.
7. W. D. Huse *et al.*, *Science* **246**, 1275 (1989).
8. D. R. Burton, *Trends Biotechnol.* **9**, 169 (1991).
9. A. S. Kang, T. M. Jones, D. R. Burton, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11120 (1991).
10. J. K. Scott and G. P. Smith, *Science* **249**, 386 (1990).
11. S. E. Cwirla, E. A. Peters, R. W. Barrett, W. J. Dower, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6378 (1990).
12. J. McCafferty, A. D. Griffiths, G. Winter, D. J. Chiswell, *Nature* **348**, 552 (1990).
13. A. S. Kang, C. F. Barbas, III, K. D. Janda, S. J. Benkovic, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4363 (1991).
14. C. F. Barbas, III, A. S. Kang, R. A. Lerner, S. J. Benkovic, *ibid.*, p. 7978.
15. G. Winter and C. Milstein, *Nature* **349**, 293 (1991).
16. C. F. Barbas, III, J. D. Bain, D. M. Hoekstra, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4457 (1992).
17. C. Chothia and A. M. Lesk, *J. Mol. Biol.* **196**, 901 (1987).
18. H. Gram *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3576 (1992).