Monoclonal Antibodies from Combinatorial Libraries

DENNIS R. BURTON

Departments of Immunology and Molecular Biology, Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received February 25, 1993

Antibody Recognition of a Vast Set of Molecular Shapes: How Nature Does It

Antibodies are molecules of immune defense with the allotted task of binding to foreign materials such as bacteria and viruses and triggering their elimination.[1-6] With no foreknowledge of the shapes to be encountered during a lifetime and given the propensity of some pathogens to rapidly change their shapes (e.g. HIV), nature has evolved a "catch-all" strategy followed by positive feedback.

The strategy operates as follows. Antibodies are made by white blood cells known as B cells, and a human being has on the order of 10^{10} such cells. Each individual cell or set of identical cells ("clones") makes multiple copies of one antibody. With reasonable assumptions about the nature of B cells,[7] it can be estimated that a human should make on the order of 10^7–10^9 different antibody molecules before any contact with a foreign material or antigen is made. This naive repertoire is expected to generate enough different molecular shapes in the antibody combining site to recognize, albeit with moderate affinity, virtually any foreign material. The naive repertoire is expressed on the surface of B cells, with each individual cell expressing many molecules of an individual antibody. When an antigen is encountered, it binds to a small number of cells which express an antibody with affinity for the antigen. These cells are stimulated and divide to produce more antigen-specific cells. During this process, a greatly increased rate of mutation occurs in the antibody genes, leading to the generation of many variants of the first selected clones. The antigen now binds to and stimulates most effectively those clones having the highest affinity for the antigen. Rounds of mutation and selection thus act to improve on the initial blueprint. At the end of the process, some cells become dedicated to secreting large amounts of high-affinity antibodies and others are retained as memory cells. The latter are immediately stimulated to produce high-affinity antibodies on renewed contact with the antigen and are responsible, at least in part, for an individual becoming "immune" to a given antigen, e.g. a virus.

A naive repertoire of 10^7 antibody molecules requires 10^7 genes, which is a considerable genetic load. To avoid carrying this number of genes in every cell of the body (the germ line), a far smaller set of gene fragments is carried, and these are randomly recombined in antibody-producing cells to generate the larger set. Each antibody molecule contains a heavy and a light chain (Figure 1). The heavy-chain gene (or more precisely the part involved in antigen binding) is formed by recombination of a V_H gene with a D gene and a J gene. As there are believed to be roughly 100 V_H genes, 30 D genes, and 6 J_H genes, this gives 18 000 heavy-chain genes. In fact, because of imprecise joining of gene segments, this number can be far higher. For the light chain, about 100 V_L genes with 5 J genes gives at least 500 combinations. Heavy–light chain combination then generates at least $18\,000 \times 500 = 9 \times 10^6$ antibody molecules and probably many more.[8] Because of imprecise joining and associated modifications, the D

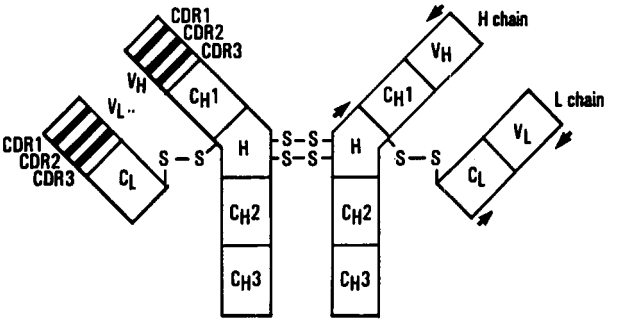

Figure 1. The antibody molecule. The human IgG1 molecule is shown schematically. Antigen binding is a combined property of the V_H (variable heavy) and V_L (variable light) domains principally, although not exclusively, through interaction with the antigen of the CDRs (complementarity determining regions) or hypervariable loops. The domain structure means that the antigen-binding properties of Fab ((V_H C_H1) (V_L C_L)) and Fv (V_H V_L) fragments are generally considered for many purposes to be equivalent to those of the whole antibody. C_H1 represents the constant heavy chain domain 1, C_L the constant light chain domain, and H the hinge region. Somewhat simplified, a V_H domain can be seen as the result of the genetic recombination of three segments. A V_H gene encodes most of the domain, a D segment encodes most of the CDR3, and a J segment encodes a part of the CDR3 and the rest of the domain. A V_L domain is the result of recombination between a V_L gene, which encodes most of the domain, and a J segment, which encodes part of the CDR3 and the rest of the domain. PCR primers are chosen to amplify DNA corresponding to the Fd part (V_H C_H1) of the heavy chain and the whole of the light chain. Equivalent positions on the protein are represented by arrows in the figure. The 3′ primer for the heavy chain hybridizes to a hinge region sequence to include the cys involved in the heavy–light chain disulfide bridge in IgG1. Other antibody isotypes have been amplified by a similar strategy (e.g., ref 45).

Dennis R. Burton was born in Preston, England, in 1952. He obtained a B.A. in Chemistry from the University of Oxford in 1974 and a Ph.D. based on nuclear magnetic resonance studies of proteins from Sture Forsén's laboratory in Lund, Sweden, in 1978. After a postdoctoral fellowship with Raymond Dwek in Oxford, he was appointed Lecturer in Biochemistry at the University of Sheffield in 1981. He was appointed Senior Lecturer in 1987 and to a Personal Chair in 1990. Since 1991, he has been a Member at the Scripps Research Institute. His research interests have focused on diverse aspects of structure–function in the antibody molecule.

(1) Stryer, L. *Biochemistry*, 3rd ed.; W. H. Freeman: New York, 1988.
(2) Watson, J. D.; Gilman, M.; Witkowski, J.; Zoller, M. *Recombinant DNA*, 2nd ed.; Scientific American Books, W. H. Freeman: New York, 1992; Chapter 16.
(3) Golub, E. S.; Green, D. R. *Immunology. A synthesis*, 2nd ed.; Sinauer: Sunderland, MA, 1991.

0001-4842/93/0126-0405$04.00/0

region is the area of greatest diversity. It corresponds to a region on the antibody molecule of great importance for antigen recognition, complementarity-determining region 3 (CDR3) of the heavy chain (Figure 1).

The antibody is then nature's most powerful system for the recognition of diverse molecular species. To understand this system and to harness its potential for medicine and for chemistry requires that one can prepare antibodies of a single defined molecular specificity, or monoclonal antibodies. This is not a trivial task given the complex mixture of antibodies present in animal serum. A great advance was made in 1975 when Köhler and Milstein[50] described the development of hybridoma technology. In this method, a single antibody-producing cell is fused with a cancer cell to generate a hybrid cell line or hybridoma which secretes a single antibody species and can be grown indefinitely. The method was applied with great success to the generation of rodent monoclonal antibodies for many applications including diagnostics, the discovery and characterization of cell surface molecules, and the mapping of protein functions. The major limitations of the method are that it has not been successful for human antibodies and that the manipulations of eukaryotic cells involved are time consuming and labor intensive. The latter greatly restricts the extent to which detailed molecular studies can be attempted. In comparison to eukaryotic cells, bacteria have many advantages and are the basis for much of modern molecular biology. The idea therefore developed that if one could express antibody repertoires in bacteria then one could truly begin to harness the molecular recognition capabilities of the antibody molecule.

Establishing Antibody Repertoires in Bacteria: Combinatorial Libraries on Phage

The demonstration that single antigen-binding fragments of antibodies (Fab, Fv) could be expressed in bacteria was an important step[9,10] in this area. Further, it became possible, using the polymerase chain reaction (PCR),[11-13] to obtain enough heavy- and light-chain DNA for cloning from a mixture of antibody-producing cells. The next stage was to choose a cloning vehicle, or vector, to get the antibody genes into bacteria and expressed. Phage, bacteria-infecting viruses, were favored because of their high efficiency of infection which allows the cloning of a large number of antibody genes. The ensemble of phage, each containing one heavy- and one light-chain gene, is described as a combinatorial antibody library. More correctly, the ensemble is a random combinatorial library since heavy- and light-chain DNA are prepared and cloned into the vector separately and no knowledge of the in vivo combination remains. The first combinatorial antibody library was constructed from an immunized mouse using λ phage as the cloning vector.[14] An array of antibody Fab fragments binding the immunizing antigen was successfully isolated from the library by probing with the antigen.

The methodology was then advanced to the expression of antibody fragment libraries on the surface of filamentous phage[15-18] building on the pioneering work of Smith on the expression of peptides at the phage surface.[47] Antibody display phage have antibody genes contained inside the phage particle and antibodies expressed on the outside. The advantage is that specific antibody–phage can now be selected from the library by binding to an immobilized antigen, a technique referred to as panning. Once selected, individual antibody–phage can be used to infect bacteria, multiply, and produce large amounts of the antibody. Thus recognition and replication become linked as for the B cell in vivo. We shall now describe the combinatorial library system in use at the Scripps Research Institute[17,19-29] in overview and then in some detail. An alternative but similar system has been developed at the Laboratory of Molecular Biology in Cambridge.[18,30-33]

Construction and Selection of Binders from Combinatorial Phage Display Libraries

An outline of the strategy is shown in Figure 2. The antibody genes are derived either from an animal or

(4) Roitt, I. M.; Brostoff, J.; Male, D. K. *Immunology*, 2nd ed.; Gower Medical Publishing: London, 1988.
(5) Harlow, E.; Lane, D. *Antibodies, A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1988.
(6) Burton, D. R. *Trends Biochem. Sci.* **1990**, *15*, 64–69.
(7) Berek, C.; Milstein, C. *Immunol. Rev.* **1988**, *105*, 5–26.
(8) Hillson, J. L.; Perlmutter, R. M. *Int. Rev. Immunol.* **1990**, *5*, 215–229.
(9) Skerra, A.; Pluckthun, A. *Science* **1988**, *240*, 1038–1041.
(10) Better, M.; Chang, C. P.; Robinson, R. R.; Horwitz, A. H. *Science* **1988**, *240*, 1041–1043.
(11) Orlandi, R.; Gussow, D. H.; Jones, P. T.; Winter, G. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3833–3837.
(12) Larrick, J. W.; Danielsson, L.; Brenner, C. A.; Abrahamson, M.; Fry, K. E.; Borrebaeck, C. A. K. *Biochem. Biophys. Res. Commun.* **1989**, *160*, 1250–1256.
(13) Sastry, L.; Alting-Mees, M.; Huse, W. D.; Short, J. M.; Sorge, J. A.; Hay, B. N.; Janda, K. D.; Benkovic, S. J.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5728–5732.
(14) Huse, W. D.; Sastry, L.; Iverson, S. A.; Kang, A. S.; Alting-Mees, M.; Burton, D. R.; Benkovic, S. J.; Lerner, R. A. *Science* **1989**, *246*, 1275–1281.
(15) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. *Nature* **1990**, *348*, 552–554.
(16) Kang, A. S.; Barbas, C. F., III; Janda, K. D.; Benkovic, S. J.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4363–4366.
(17) Barbas, C. F., III; Kang, A. S.; Lerner, R. A.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7978–7982.
(18) Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. *Nature* **1991**, *352*, 624–628.
(19) Burton, D. R.; Barbas, C. F., III; Persson, M. A. A.; Koenig, S.; Chanock, R. M.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10134–10137.
(20) Barbas, C. F., III; Björling, E.; Chiodi, F.; Dunlop, N.; Cababa, D.; Jones, T. M.; Zebedee, S. L.; Persson, M. A. A.; Nara, P. L.; Norrby, E.; Burton, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9339–9343.
(21) Barbas, C. F., III; Crowe, J. E., Jr.; Cababa, D.; Jones, T. M.; Zebedee, S. L.; Murphy, B. R.; Chanock, R. M.; Burton, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10164–10168.
(22) Zebedee, S. L.; Barbas, C. F., III; Hom, Y.-L.; Caothien, R. H.; Graff, R.; DeGraw, J.; Pyati, J.; LaPolla, R.; Burton, D. R.; Lerner, R. A.; Thornton, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3175–3179.
(23) Duchosal, M. A.; Eming, S.; Fischer, P.; Leturcq, D.; Barbas, C. F., III; McConahey, P. J.; Caothien, R. H.; Thornton, G. B.; Dixon, F. J.; Burton, D. R. *Nature* **1992**, *355*, 258–262.
(24) Gram, H.; Marconi, L.-A.; Barbas, C. F., III; Collet, T. A.; Lerner, R. A.; Kang, A. S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3576–3580.
(25) Barbas, C. F., III; Bain, J. D.; Hoekstra, D. M.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4457–4461.
(26) Williamson, R. A.; Burioni, R.; Sanna, P.; Partridge, L. J.; Barbas, C. F., III; Burton, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4141–4145.
(27) Kang, A. S.; Jones, T. M.; Burton, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11120–11123.
(28) Barbas, C. F., III; Collet, T. A.; Amberg, W.; Roben, P.; Binley, J. M.; Hoekstra, D.; Cababa, D.; Jones, T. M.; Williamson, R. A.; Pilkington, G. R.; Haigwood, N. L.; Satterthwait, A. C.; Sanz, I.; Burton, D. R. *J. Mol. Biol.* **1993**, *230*, 812–823.
(29) Collet, T.; Roben, P.; O'Kennedy, R.; Barbas, C. F., III; Burton, D. R.; Lerner, R. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10026–10030.
(30) Hoogenboom, H. R.; Griffiths, A. D.; Johnson, K. S.; Chiswell, D. J.; Hudson, P.; Winter, G. *Nucleic Acids Res.* **1991**, *19*, 4133–4137.
(31) Marks, J. D.; Hoogenboom, H. R.; Bonnert, T. P.; McCafferty, J.; Griffiths, A. D.; Winter, G. *J. Mol. Biol.* **1991**, *222*, 581–597.

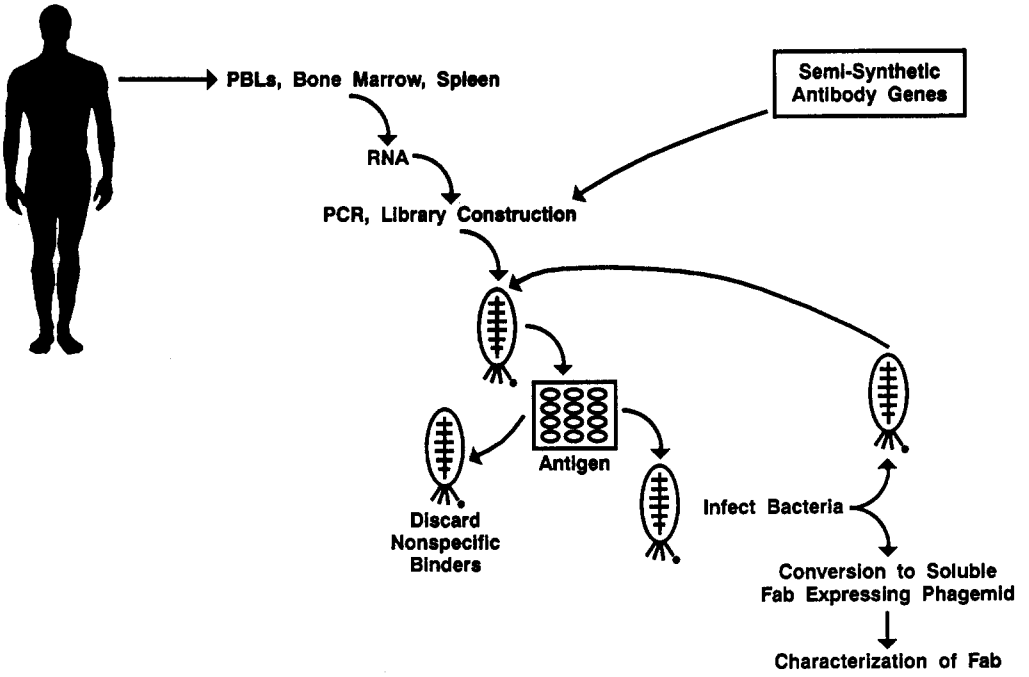

Figure 2. Strategy for cloning monoclonal Fab fragments from combinatorial libraries on the surface of phage. RNA is prepared from a tissue source containing antibody producing cells, e.g. from peripheral blood, bone marrow, or the spleen of a human. The mRNA is reverse transcribed to cDNA, and then the Fd part of the heavy chain and the light chain are amplified using the PCR reaction (see caption to Figure 1). Alternatively, heavy- and light-chain DNA are amplified from semisynthetic genes. The PCR primers incorporate restriction sites which can now be used to clone the PCR products into a phagemid vector (Figure 3). The ensemble of heavy-chain inserts in the vector is referred to as a heavy-chain library, and similarly for the light chain. When both chains are present, the ensemble is referred to as a combinatorial (Fab) library. The combinatorial phagemid library is now "rescued" to a phage display library (Figure 4) in which each phage expresses an Fab on the surface and has the corresponding genes inside the phage. This library is now "panned" against the immobilized antigen (Figure 5), and specific phage–Fabs are selected by their ability to bind to the antigen. These specific phage–Fabs can now be converted to a phagemid form which, when inserted into bacterial cells, produce soluble Fabs. This figure is adapted from ref 52.

human (immunized or nonimmunized), or semisynthetically. In the former case, antibody-producing cells are isolated and RNA is prepared and reverse transcribed into DNA in a reaction catalyzed by the enzyme reverse transcriptase. Heavy (Fd part) and light chains are then amplified using the PCR reaction as described in the caption to Figure 1. This provides the genetic information necessary to produce the Fab antigen-binding fragment of the antibody molecule. The heavy- and light-chain DNA are then cut with restriction enzymes and cloned sequentially into a phagemid vector designated pComb3,[17] as depicted in Figure 3. This vector is then inserted into bacterial cells (the cells are said to be "transformed") by the technique of electroporation, which renders the cell membranes permeable to DNA. One phagemid, carrying one set of antibody genes, enters each cell. There it is capable of producing antibody Fab fragments which assemble in the periplasm of the cell anchored in the inner membrane (Figure 3). The efficiency of the transformation process with this vector is on the order of 10^8 per μg of DNA. It is this step that limits library size to on the order of 10^8–10^9. Because about 200 ng of DNA is used in a typical transformation, a larger size would require an extremely large number of electroporation procedures. Bacterial cells carrying plasmid are now selected by growth in a medium containing ampicillin: only the phagemid, and not the bacterial chromosome, carries a gene for ampicillin resistance (β-lactamase).

The next stage is the rescue of the phagemid DNA by adding infective helper phage to the cells. In a complex process, depicted in Figure 4, the helper phage enters the cell and packages the DNA from the phagemid (carrying the antibody genes) in preference to its own. As it leaves the cell, the phage is capped at one end by 3–5 molecules of a protein known as coat protein III (cpIII). Some of the cpIII molecules are linked to Fab molecules so a typical phage will carry one Fab molecule on its surface and the corresponding antibody genes inside. This rescue process applied to a library of phagemids generates a library of phage–Fabs. The process also serves to amplify the initial library so that typically 10^5 copies of each phage (and therefore each set of antibody genes) are generated.

This library is now "panned" against the antigen as shown in Figure 5. Typically the library will contain 10^{12} phage–Fabs (about 10^5 copies of each member of the initial library as above) in about 50 μL of solution. This solution is layered onto an antigen immobilized on plastic and incubated, and then nonspecifically bound phage are removed by repeated washings. The remaining phage (greatly enriched for specific phage–Fabs) are then eluted by, for example, acid or excess soluble antigen. This process, i.e., a round of panning, generally produces an enrichment factor for specific phage on the order of 100. The panning is then repeated, usually 3 or 4 times, which should be sufficient to extract specific phage occurring only singly in the initial library, i.e. with frequencies on the order of 1 in 10^7. Finally, the specific phage–Fabs are converted to

(32) Marks, J. D.; Griffiths, A. D.; Malmqvist, M.; Clackson, T. P.; Bye, J. M.; Winter, G. *Bio/Technology* **1992**, *10*, 779–783.
(33) Hoogenboom, H. R.; Winter, G. *J. Mol. Biol.* **1992**, *227*, 381–388.

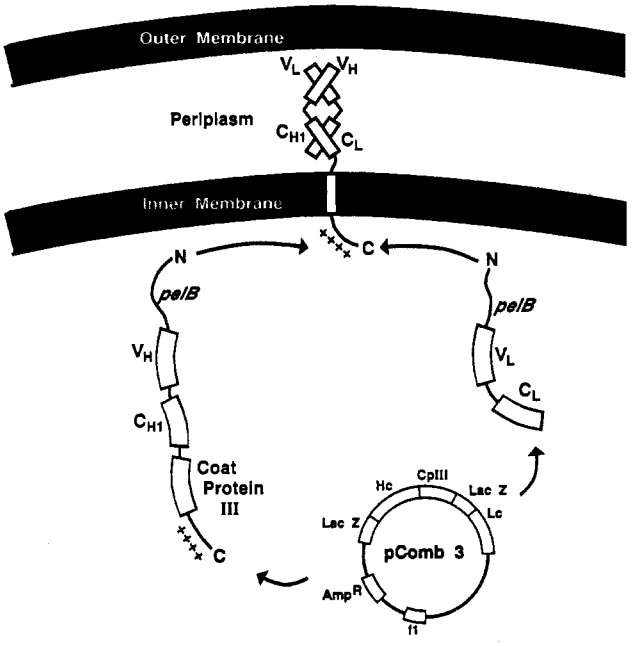

Figure 3. The composition of the pComb3 vector and the proposed pathway for Fab assembly. Expression of the Fd (Hc, heavy chain)/cpIII fusion and the light chain (Lc) is controlled by Lac promoter/operator sequences. The chains are directed to the periplasmic space by pelB signal sequences which are subsequently cleaved. The heavy chain is anchored to the membrane by the cpIII fusion, whereas the light chain is secreted into the periplasm. The two chains then assemble on the membrane. This figure is adapted from ref 25. Filamentous phage biology is reviewed in ref 46, and phage display was first described in ref 47. Comparable vector systems have been described in other laboratories, e.g. see refs 44 and 48.

plasmids expressing soluble Fabs by excision of the cpIII gene (which otherwise directs the Fab to the phage surface) and religation.

The plasmids are used to transform bacteria, and individual clones, carrying one plasmid and one set of antibody genes, grown up. The periplasmic cell extracts from each clone contain Fab, which is now screened for reactivity with the antigen, usually in a conventional ELISA assay.

Antibodies Derived from Immune Donors

Antigen-specific Fab fragments have been derived from libraries prepared from immunized mice and humans. In principle, any animal can be used, provided enough sequence information is available to design PCR primers capable of amplifying a substantial part of the antibody repertoire. Specific Fab fragments have also been prepared from humans who have not been actively immunized but have had contact with an antigen, most usually an infectious disease, at some point in time.[26] This could be many years prior to making the library. Probably the most useful prognostic indicator for the library approach is a titer of the antibody against the antigen in the serum of the donor animal. In other words, if a reasonable level of specific polyclonal antibodies to a given antigen can be identified in the serum of the donor animal, then it is likely that a library prepared from the appropriate donor tissue will yield monoclonal antibodies to the antigen. It should also be noted that Fab fragments specific for many different antigens can be isolated from the same library. Thus, if for example, mouse antibodies were required against a set of different antigens, it would not be necessary to immunize individual mice. All the antigens could be introduced into a single mouse, one library prepared, and selection introduced during the panning process.

The affinities, for the relevant antigen, of Fab fragments isolated from immune libraries appear to be in the range 10^7–10^9 M^{-1}.[17-21] Many of the affinities for binding to proteins are estimates from inhibition ELISA studies, although surface plasmon resonance[34] studies confirm these values. These affinities are similar to those found for monoclonal antibodies derived from application of the hybridoma approach to immunized mice.[5,35]

The number of different antibodies generated against the immunizing antigen can be readily assessed by sequencing the antibody genes on the appropriate plasmid DNA. Typically one finds a limited number of heavy chains combined with a greater number of light chains.[18,20-23,26-27,36] In other words, a given heavy chain is often found in combination with several different light chains with retention of antigen affinity. This phenomenon is referred to as chain promiscuity. Frequently, families of closely related sequences are seen. As an example, 33 Fabs reacting with high affinity with the HIV-1 surface glycoprotein gp120 were obtained from a library derived from an HIV-1 seropositive individual.[28] Seven groups of heavy chains were described in which the CDR3 regions were identical, or nearly so, implying that the chains arose from a common precursor (i.e. the same VDJ gene rearrangement). In some groups, the heavy chains were paired with closely similar light chains. In others, they were paired with very different light chains. Simply by panning the library, therefore, one derives families of related antibodies. These can be further manipulated as discussed below.

Genetic Manipulation of Fabs

The cloning of Fabs in bacteria means that, following the initial identification of antigen-specific Fabs, the clones may be manipulated to improve their affinity or alter their specificity. One strategy, "chain shuffling", should prove to be of general use in altering the properties of antibodies isolated from combinatorial libraries. We summarize here our own experiences from two approaches to chain shuffling using a set of HIV-1 gp120-binding Fabs.

In the first, a given chain was recombined with the complete original library of the complementary chain.[28] We have found that this approach, with either chain fixed, reveals a spectrum of related clones, all of comparable affinity for the antigen as in the original pairing.

In the second approach to shuffling, heavy and light chains were cloned into separate plasmid vectors with differing antibiotic resistances.[29] Cells were then transformed, and double transformants expressing both chains were selected. This system revealed, in a set of 21 Fabs binding to gp120, a surprising degree of chain

(34) Jönsson, U.; Malmqvist, M. In *Advances in Biosensors*; Turner, Ed.; JAI Press: San Diego, CA, 1992; pp 291–336.
(35) Mason, D. W.; Williams, A. F. In *Handbook of Experimental Immunology*, 4th ed.; Weir, D. M., Ed.; Blackwell: Oxford, 1986; Vol. 1, Immunochemistry, Chapter 38.
(36) Persson, M. A. A.; Caothien, R. H.; Burton, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2432–2436.

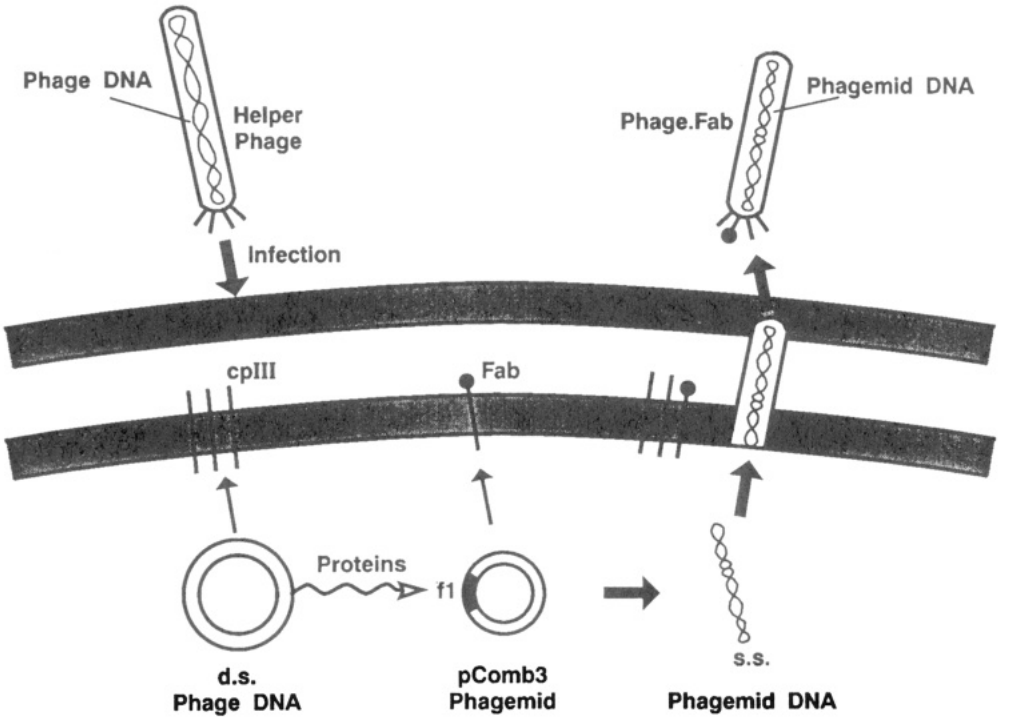

Figure 4. Helper phage rescue of phagemids to give a phage display library. Helper phage, a single-stranded DNA virus in a protein coat, infects *E. coli*, and the single-stranded DNA is converted to double-stranded DNA. This codes for a number of phage proteins. Some of these are coat proteins which accumulate in the inner membrane of the cell. Others act on the pComb3 phagemid DNA, because of the f1 sequence (Figure 3), causing the packaging of a single-stranded copy of the phagemid DNA. This is packaged in preference to the helper phage DNA which has been mutated to decrease packaging efficiency. As the assembling phage is extruded from the cell, it is "capped" by coat protein III which includes native cpIII and the cpIII–Fab complex. Thus a phage ends up with an Fab displayed on the surface and the phagemid DNA (containing the corresponding antibody genes) inside. The advantage of phagemid rescue as opposed to directly cloning antibody genes into the phage are 2-fold. Direct cloning would mean that every cpIII molecule would carry an Fab. This would greatly reduce infectivity since cpIII is involved in entry to *E. coli*. Furthermore, multivalent display is expected to hinder selection of binders on the basis of affinity in the panning process because of chelation effects. Phagemid rescue probably leads to primarily monovalent Fab display.

promiscuity, particularly in the heavy chain. Thus some heavy chains were found which would productively pair with all light chains in the set. Other heavy chains were more selective in their partners. The system was further used to look at recombination of chains from Fabs binding to gp120 with those binding to tetanus toxoid. The results clearly supported the notion that specificity for the antigen is dominated by the heavy chain. Heavy chains from some gp120 binders could retain affinity for gp120 with a light chain from a tetanus toxoid binder. A heavy chain from a tetanus toxoid binder could retain affinity for toxoid with light chains from gp120 binders. However, none of the light chains in the experiment could dictate a new specificity to heavy-chain partners.

One of the consequences of this extensive promiscuity is that it may be possible to introduce cofactors attached to the light chain. For instance, one could envisage immunizing with a given transition-state analog, deriving binders, and then shuffling functional heavy chains against a library of metallo–light chains to access catalytic antibodies.[53]

Alternative strategies for the refinement of clones involve mutagenesis and reselection. Random mutagenesis of the variable regions is possible by a number of different approaches such as chemical mutagenesis,[37] polymerase-induced mutagenesis,[38] and in vivo mutagenesis using mutator strains of *E. coli*.[39] Focused mutagenesis, in which several residues are targeted, does not mimic the supposed random mutation and selection of the immune system, but it does allow all possible mutations in a defined region to be explored. This strategy has been successful in generating high-affinity variants of the human growth hormone[40,41] and, when applied to the CDRs of antibodies, in improving affinities for antigens.[25,33]

Although positive selection for variants of increased affinity is an obvious aim, mutagenesis can also be employed to increase or decrease cross-reactivity.[42] The panning process can assist here. For instance, specificity can be increased by including the antigen that gives rise to unwanted cross-reactivity in the wash solution during the selection or by preselection of the phage with the antigen. Conversely, cross-reactivity can be encouraged by panning alternately against the antigens in question.

Another genetic manipulation of particular importance for any therapeutic application of library-derived antibodies is to splice the Fabs to Fc (Figure 1) to generate whole antibody molecules. This is relatively straightforward. Expression of whole antibody mol-

(37) Myers, R. M.; Lerman, L. S.; Maniatis, T. *Science* **1985**, *229*, 242–247.
(38) Leung, D. W.; Chen, E.; Goeddel, D. V. *Technique* **1989**, *1*, 11–15.
(39) Schaaper, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8126–8130.
(40) Bass, S.; Greene, R.; Wells, J. A. *Proteins: Struct., Funct., Genet.* **1990**, *8*, 309–314.
(41) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. *Biochemistry* **1991**, *30*, 10832–10838.
(42) Glaser, S. M.; Yelton, D. E.; Huse, W. D. *J. Immunol.* **1992**, *12*, 3903–3913.

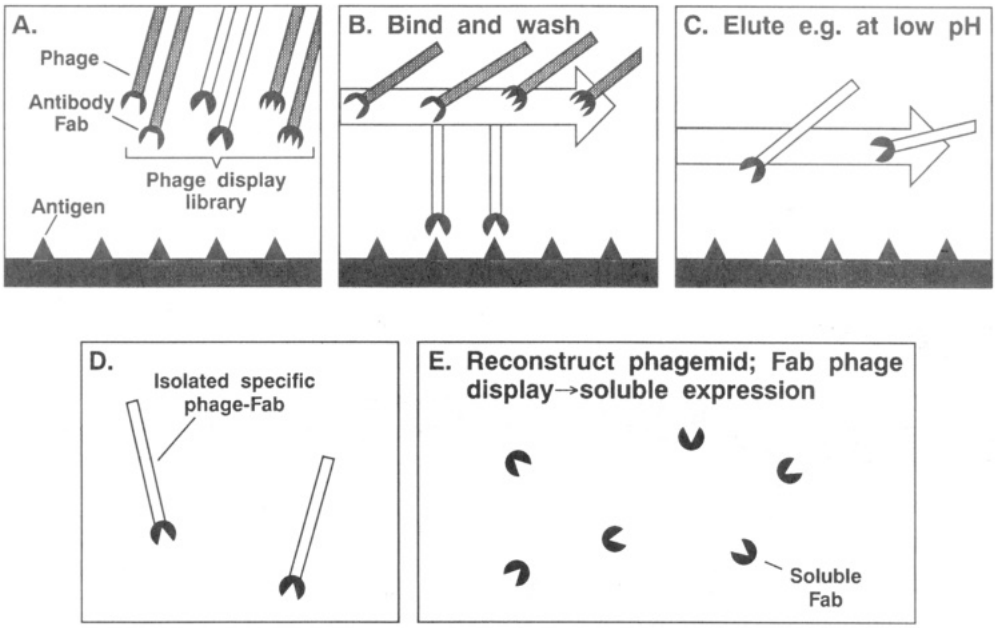

Figure 5. Panning for selection of specific Fabs from the combinatorial phage display library. The antigen (immobilized for example on a plastic surface) is exposed to the phage display library (A) when specific phage–Fabs bind, and most of the rest of the phage are removed by washing (B). Specific phage–Fabs are then eluted by low pH or excess soluble antigen (C) and isolated (D). In practice, after one round of panning, many irrelevant phage–Fabs are present at D and the process A–D is repeated with several rounds of panning. After each round of panning, the eluted phage–Fabs are amplified to minimize the chances of losing important phage. Amplification is carried out by infecting bacterial cells with the phage (which are then converted from single-stranded phage to double-stranded phagemid DNA), growing up the cells, and then rescuing phage–Fabs using helper phage as in Figure 4. Finally the phage–Fabs from D are converted to the phagemid form, the DNA is prepared, and the gene for cpIII is excised. Religation then gives a reconstructed phagemid which can be used to transform bacterial cells for the production of soluble Fab fragments. This figure was modeled after ref 49.

ecules, at least in part because of the glycosylation requirement of the Fc part of the molecule, is then carried out in eukaryotic cell lines. We have shown how a whole IgG1 molecule can be expressed in Chinese hamster ovary (CHO) cells using the Fd and light chains derived from phages.[51] As expected, the molecule binds the antigen with retained antigen affinity.

Antibodies Derived without Immunization

There are two principle types of libraries for derivation of antibodies without immunization. The first is termed a "naive" library, which is derived from an animal but typically from the IgD or IgM class of antibodies rather than the more usual IgG class. The former should be less influenced by the immunization history of the animal. The antibodies selected by this route[24,31,32] will tend to be of lower affinity than those that have been subjected to rounds of mutation and selection in vivo. Nevertheless, by subjecting the first selected antibodies to some of the procedures described above, high-affinity antibodies can be generated. For instance,[32] an antibody with an affinity of 3×10^6 M^{-1} for a small hapten was selected from a nonimmunized donor. Light-chain shuffling increased this value by 20-fold, and shuffling of the CDR1 and 2 of the heavy chain caused a further 15-fold increase to 1×10^9 M^{-1}.

The second type of library is referred to as semisynthetic in that the antibody frameworks are derived from natural antibodies but some or all of the CDRs are derived from synthetic gene segments. Initial explorations of semisynthetic antibodies utilized a single clone selected from a library from an immune donor.[25] A 16 amino acid random sequence was then introduced over the CDR3 region of the heavy chain to generate a vast library of antibodies. Selection of the library against a variety of antigens allowed for the cloning of new specificities. The complete randomization of 16 amino acids would require the generation of a library of greater than 10^{20} clones, far in excess of the number that is obtainable by the transformation of *E. coli*. However, libraries can be constructed that match or exceed the diversity of clones examined by an animal at a given moment, approximately 10^7. Selection of this library derived from a human anti-tetanus clone against a new antigen, fluorescein, resulted in the isolation of clones with affinities for fluorescein approaching that obtained by the secondary boost of a mouse. This strategy is proving useful for the generation of antibodies against a variety of antigens and also has been demonstrated in the Cambridge laboratory.[33] Extension of this strategy to the synthesis of all the CDRs or the use of natural libraries of FR1–FR3 fragments in combination with synthetic CDR3s should yield libraries from which almost any given specificity is retrievable. This approach has one distinct advantage over naive libraries. Diversity of these libraries is controlled at the level of nucleic acid synthesis, whereas the diversity of a naive library is limited by the source of RNA, which is susceptible to bias by RNA derived from plasma cells or activated B cells.

The generation of antibodies without immunization[43,44] clearly has many advantages over the use of immunized animals. There are two principal disadvantages currently. One is that a greater diversity of specific antibodies is typically obtained from an im-

(43) Lerner, R. A.; Kang, A. S.; Bain, J. D.; Burton, D. R.; Barbas, C. F. *Science* **1992**, *258*, 1313–1314.
(44) Marks, J. D.; Hoogenboom, H.; Griffiths, A. D.; Winter, G. *J. Biol. Chem.* **1992**, *267*, 16007–16010.

munized library, although this is liable to change as more semisynthetic libraries are produced. The second is that, for human therapy, the extent to which semisynthetic antibodies will be perceived as foreign by the human is unknown. Nevertheless semisynthetic libraries in particular offer a myriad of new possibilities as discussed below.

Future Prospects

Combinatorial libraries offer ready access to human monoclonal antibodies from immune donors. Therefore we can expect a thorough evaluation of the use of human antibodies in the prophylaxis and therapy of infectious diseases, particularly viral diseases where current drugs have limited efficacy. In bacterial diseases, the ability to neutralize toxins may be of greater importance. For applications to cancer or as agents to modulate immune responses, the relevant antibodies would need to be directed to self-antigens and immune donors may be more difficult or impossible to find. Semisynthetic libraries and in vitro gene manipulation may then be crucial. Human Fab fragments produced in bacteria may prove attractive diagnostic reagents because of their cost. The library approach is also likely to stimulate interest in studying antibody responses in humans although the loss of in vivo chain pairing should be noted. The understanding of autoimmunity and the precision of vaccine design should benefit from such study of antibody responses.

For those interested in molecular recognition, the most exciting developments are likely to come from semisynthetic libraries. The combination of randomization of antibody sequences known generally to be involved in recognition (the CDRs) and efficient selection by phage display should provide interesting insights into macromolecular interactions. The approach has advantages over site-directed mutagenesis in that it makes fewer assumptions about our understanding of protein structure–function. Nevertheless, some element of design may be beneficially coupled to randomization and selection. This is apparent in the generation of high-affinity antibodies to human cell receptors (C. F. Barbas, personal communication) where a core receptor binding motif was incorporated into one of the CDRs and adjacent sequences were randomized prior to selection. A blend of design and randomization followed by selection may also ultimately prove to be valuable in persuading antibodies to catalyze chemical reactions in the catalytic antibody field. Finally, given recent developments, the appearance of a machine to generate and evolve antibodies is a real possibility.

I acknowledge the advice and help of many valued colleagues but particularly Richard Lerner and Carlos Barbas at the Scripps Research Institute. Some of the work described is supported by the NIH and by Johnson and Johnson.

(45) Walker, M. R.; Bevan, L. J.; Daniels, J.; Rottier, M. M. A.; Rapley, R.; Roberts, A. M. *J. Immunol. Methods* **1992**, *149*, 77–85.
(46) Model, P.; Russel, M. *The Bacteriophages*; Calendar, R., Ed.; Plenum Publishing Corporation: New York, 1988; Vol. 2, pp 375–456.
(47) Smith, G. P. *Science* **1985**, *228*, 1315–1316.
(48) Garrard, L. J.; Yang, Y.; O'Connell, M. P.; Kelley, R. F.; Henner, D. J. *Bio/Technology* **1991**, *9*, 1373–1377.
(49) Burton, D. R. *Hosp. Pract.* **1992**, *27*, 67–74.
(50) Köhler, G.; Milstein, C. *Nature* **1975**, *256*, 495–497.
(51) Bender, E.; Woof, J. M.; Atkin, J. D.; Barker, M. D.; Bebbington, C. R.; Burton, D. R. *Hum. Antibod. Hybridomas* **1993**, *4*, 74–79.
(52) Burton, D. R.; Barbas, C. F., III. *Chem. Immunol.* **1993**, *56*, 112–126.
(53) Sarvetnick, N.; Gurushanthaiah, D.; Han, N.; Prudent, J.; Schultz, P.; Lerner, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4008–4011.