Intrinsic and Extrinsic Factors in Protein Antigenic Structure

Jay A. Berzofsky

Summary. Recent advances in the preparation of synthetic peptide vaccines and the use of synthetic peptides as probes of antigenic structure and function have led to renewed interest in the prediction of antigenic sites recognized by antibodies and T cells. This review focuses on antibodies. Features intrinsic to the antigen, such as hydrophilicity and mobility, may be useful in the selection of amino acid sequences of the native protein that will elicit antibodies cross-reacting with peptides, or sequences which, as peptides, will be more likely to elicit antibodies cross-reactive with the native protein. Structural mobility may also contribute to protein-protein interactions in general. However, the entire accessible surface of a protein is likely to be detectable by a large enough panel of antibodies. Which of these antibodies are made in any individual depends on factors extrinsic to the antigen molecule, host factors such as self-tolerance, immune response genes, idiotype networks, and the immunoglobulin structural gene repertoire.

Interest in defining and predicting antigenic sites of proteins stems, in part, from the recent advances in the production of synthetic vaccines (*1*, *2*) and from the development of techniques for the cloning and sequencing of genes. Even in the absence of a known gene product, synthetic peptides can be made on the basis of the DNA sequence, and antibodies to these peptides can be used to isolate and characterize the unknown gene product (*1*). This interest goes back to Landsteiner (*3*), who studied antigen structure before anything was known of the structure of an antibody. A fundamental problem is that, experimentally, an antigenic site (also called an antigenic determinant or epitope) can *only* be defined by examining the products of the immune response (antibodies or lymphocytes) of a particular animal or person (the host) that has been exposed to the antigen. Thus, for example, the portion of a protein bound by a specific antibody molecule can be called the antigenic site recognized by that antibody. The fundamental question is, can that site be called an antigenic site in its own right, or only with respect to the particular antibody which binds that site, or with respect to a host that can make such antibodies.

Dr. Berzofsky is a Senior Investigator in the Metabolism Branch of the National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205.

One viewpoint is that certain parts of proteins are inherently antigenic sites and that this property would be intrinsic to the nature of the protein molecule and independent of the host to be immunized (*4*). On the basis of this concept, attempts have been made to define the properties of certain protein substructures which might make them inherently antigenic. To design synthetic vaccines or prepare antibodies to proteins that have not been isolated and whose primary sequence is known only from the DNA sequence, it is important to be able to predict which stretch of primary sequence is most likely to elicit antibodies that will also bind to the native protein.

The alternative viewpoint is that virtually any accessible part of a protein is potentially an antigenic site, and that the choice of sites that elicit an immune response in a particular case depends largely on the bias of the immune system of a specific host (*5*). For instance, for mammalian proteins injected into a mammalian host, self-tolerance to the homologous host protein may strongly influence the result. Even in the case of a nonmammalian protein for which there is no host counterpart (such as viral, bacterial, or invertebrate parasite antigens), the regulatory mechanisms of the host, such as immune response genes linked to the major histocompatibility complex of transplantation antigens (*6*), or networks of idiotype–anti-idiotype interactions (*7*), will determine the outcome of an immunization. Such regulatory mechanisms may also involve the interplay of helper and suppressor T lymphocytes, each with its own repertoire of antigenic specificities (*6*, *8*). If this viewpoint is correct, the ability to rationally design synthetic vaccines or antibody probes would be limited and would require a more empirical approach, although comparisons of the antigen with homologous host proteins would provide a rational starting point. However, from a more optimistic point of view, if any accessible part of a protein is a potential antigenic site, the likelihood of being able to synthesize a useful site in the absence of much secondary or tertiary structural information about a protein is greatly increased.

The purpose of this review is to examine the evidence for each of these hypotheses, and to distinguish between properties intrinsic to the antigen and factors dependent on the host that determine the epitope specificity of an immune response. Although protein antigenic sites recognized by T cells have recently been characterized (*9*, *10*), this review will be limited to the regulation of antibody specificity.

First, we must consider the semantic problem of immunogenicity versus antigenicity. Immunogenicity refers to the ability to elicit an immune response (antibody or T cell) when used to immunize an animal, whereas antigenicity refers to the ability to be recognized by the product of a previous immune response, either antibody or T cell. Although these properties often coincide, a few examples may illustrate the distinction. A hapten is defined as a small molecule (such as dinitrophenol) that is not immunogenic alone but, when attached to an immunogenic protein (called the carrier), will elicit antibodies that can bind to the free hapten. Thus, the free hapten is antigenic without being immunogenic. In another situation, although hen lysozyme is not immunogenic in a strain of mice such as C57BL/10 for reasons of T-cell suppression or Ir gene control (or both), when these mice are immunized with a fragment of hen lysozyme or with certain other lysozymes, the mice can still make antibodies that will bind cross-reactively to hen lysozyme (*8*). In this case, the sites on hen lysozyme bound by these antibodies are antigenic without being immunogenic in this strain. However, by definition, anything that is immunogenic must also be antigenic.

In principle, antigenic sites can be divided into two structural categories

(*5*). A segmental (continuous) site exists wholly within a continuous segment of the amino acid sequence. An assembled topographic site consists of amino acid residues far apart in the primary sequence but brought together in the surface topography of the native protein by the way it folds in three dimensions (*5, 11, 12*). These categories are the outgrowths of the earlier categorizations of sequential and conformational sites (*13*), respectively. However, it has become clear that even segmental sites bind with highest affinity in a preferred conformation (*14, 15*). Furthermore, peptides have been synthesized either with the correct order of L-amino acids or with the reverse order of D-amino acids, which should result in approximately the same arrangement and order of amino acid side chains but a different peptide backbone configuration (*1*). Antibodies to the former did not bind the latter, an indication that recognition involves more than just a linear sequence of side chains. Finally, knowing that an antibody combining site consists of a three-dimensional array of amino acids with a defined surface contour, the prediction can be made that among various possible conformations of a polypeptide sequence, some would fit better into the antibody combining site than others and thus bind with higher affinity. In other words, the three-dimensional conformation of the antibody combining site defines an antigen conformation most complementary to itself. In this sense, all antigenic determinants must be conformational (*5, 15*).

The type of antigenic site observed depends in part on the probe or method used to investigate such sites. Early investigators of polyclonal antisera to sperm whale myoglobin used synthetic peptides or proteolytic peptides as probes and detected only segmental sites (*4, 16*). It was claimed that all of the antigenic activity of the protein could be accounted for by five short segments of the sequence consisting of six to seven residues each (a total of only about 20 to 25 percent of the sequence) (*4*), no matter which species of animal was immunized (*4*). However, with the advent of monoclonal antibodies (*17*), studies could be carried out that would not have been interpretable with heterogeneous serum antibodies. By use of native myoglobins with known sequence differences, rather than short peptides, it was found that most of the monoclonal antibodies studied reacted with assembled topographic sites, not segmental ones (*11, 12*). Moreover, these were mostly not in the region of the five segments delimited with synthetic peptides. How frequently do antibodies to assembled topographic sites occur in polyclonal antisera made against native myoglobin? A minimum estimate was obtained by exhaustive depletion of antibodies that could bind to any of the three large cyanogen-bromide cleavage fragments that together span the entire sequence of myoglobin (*18*). For each serum, from three different species (goat, sheep, and rabbit), there were always 28 to 40 percent of the antibodies remaining that could bind with high affinity to native myoglobin but which failed to bind to any of the peptide fragments in a radioimmunoassay. Even conformation-dependent antibodies to segmental sites should have been removed on the affinity columns because very low affinity binding to peptides is boosted by attachment of one of the reactants to a solid phase (*19*) and by the multiple plating efficiency of the column. Thus, more than a third and perhaps a majority of antibodies made by immunizing with native protein react with assembled topographic sites.

Similar results were found for a large number of monoclonal and polyclonal antibodies to hen lysozyme (*20*). A number of assembled topographic antigenic sites as well as some that are segmental but conformation-dependent have been defined by careful mapping, involving sequence variants (*21*), substrate or enzyme inhibitor competition, competitive binding of clusters of monoclonal antibodies, and even x-ray crystallography. Also, of the monoclonal antibodies to influenza hemagglutinin, influenza neuraminidase (*22*), and tobacco mosaic virus coat protein (*23*), many bind to assembled topographic sites.

The importance of physical contiguity can be understood by examining a space-filling model of a native globular protein such as myoglobin (Fig. 1). In contrast to a stick-figure model in which the polypeptide backbone is easy to trace, the space-filling model presents a continuous surface of abutting atoms, obscures the helices, and makes it difficult to distinguish between sequentially distant residues that are adjacent on the surface because of protein folding and those that are adjacent by being neighbors in the primary sequence. An antibody that binds to a native protein must interact with a surface more like that of the space-filling model. The probability that all of the antigenic contact residues (those that fit within the antibody combining site and contribute to the affinity of binding) happen to come from the same continuous segment of polypeptide chain should be extremely low.

A second way in which the type of probe used influences the repertoire of antibodies detected is illustrated by the use of synthetic peptides of varying length. In addition to the five segmental sites defined by short peptides (*4*), other segmental sites were identified when larger peptides of myoglobin were used (*24*) (see Fig. 1). Similarly, in the case of tobacco mosaic virus protein, synthetic peptides 13 to 20 residues in length were able to detect antibodies to regions of the molecule previously thought to be nonantigenic because of lack of antibody binding to hexapeptides within these segments (*25*). The ability to detect additional antibodies with longer peptides may be due to the involvement of more than six or seven sequential residues within the site or to greater stabilization of certain secondary structures in the longer peptides. Although peptides of 15 to 20 residues are still too short to stabilize an α helix in water, significant helix formation was detectable, in the case of the myoglobin peptides 25 to 55 and 72 to 88, when they were studied by circular dichroism in trifluoroethanol (*24*). An additional caveat regarding peptide length has come from the studies of short myoglobin synthetic peptides. The binding of shorter peptides of two to seven residues was dominated by nonspecific influences of charge and hydrophobicity, so that antibodies to staphylococcal nuclease would bind to these myoglobin peptides as well as antibodies made against myoglobin. Only when the peptide length was closer to 15 to 20 residues was biologically meaningful binding detected (*24*).

When all of these approaches are combined, it is generally found that virtually the entire accessible surface of the protein can be immunogenic in one animal or another. In addition to myoglobin, this is true for lysozyme (*20, 26*), bovine and human serum albumin (*5, 27*), tobacco mosaic virus protein (*23, 25*), and influenza neuraminidase (*22*). Even for influenza hemagglutinin, (*22*), for which only four discrete clusters of antigenic sites had been detected originally, studies with larger panels of antibodies showed the originally discrete regions of antigenicity to overlap as a continuum.

A continuum of antigenic sites over virtually the entire surface of many proteins was also indicated by pairwise comparisons among proteins of evolutionarily related species, such as lysozymes, myoglobins, ribonucleases, cytochromes c, azurins, and albumins. These studies demonstrated, by effects on cross-reactivity, that about 80 percent of the amino acid substitutions, scattered around the protein surface, were detectable by antibodies (*5, 26*).

All of these results lead to the formulation of a "multideterminant-regulatory hypothesis" (*5*). This hypothesis suggests that the surface of a protein is a continuum of potential antigenic sites, and that the selection of sites that elicit antibody production in a given animal depends on a number of regulatory influences in the host. Before discussing these extrinsic factors, we must examine the factors intrinsic to the protein antigen that have been considered as possibly predictive of antigenic sites.

Intrinsic Properties of the Antigen Used as Predictors of Antigenic Sites

A number of approaches have been taken to search for sites that are intrinsically antigenic and for principles that would predict such sites. Notwithstanding the evidence that the entire surface of a protein may be antigenic, these principles may be useful in studying relative antigenic potency, in analyzing the forces involved in antigen-antibody interactions, and in selecting optimal sequences to use as synthetic immunogens.

Accessibility. As a minimum requirement, antigenic sites should be accessible on the surface (*28*) in order to be detected by antibodies binding to the native protein. For example, studies of monoclonal antibodies to influenza neuraminidase showed the surface of the protein heads to be a continuum of overlapping antigenic sites (*22*), but no antibodies were found binding to the stalks. This apparent paradox was explained when it was noted that the protein of the stalks was almost entirely covered by carbohydrate and thus not accessible (not to say that carbohydrate cannot also be immunogenic).

Although accessibility is necessary, it is not known whether accessibility of a site on the native protein surface is sufficient for antibody binding. Also, accessibility is useful for predicting antigenic sites only for proteins whose three-dimensional structure is known from x-ray crystallography. It would be more broadly useful to have a method of predicting sites from primary sequence.

Hydrophilicity. Water-soluble globular proteins, to be stable in the aqueous environment, fold in the native conformation so as to bury hydrophobic residues in the interior and expose mostly hydrophilic ones on the surface. Hopp and Woods (*29*) suggested that possible antigenic sites might be predicted from primary sequence data by determining the most hydrophilic segments of the sequence. They averaged hydrophilicity parameters for overlapping segments of six amino acid residues, assigning to residue 1 the average for residues 1 to 6, to residue 2 the average for residues 2 to 7, and so forth. By analyzing 12 proteins for which antigenic sites had been reported, they found that the segment of greatest hydrophilicity for the entire protein was invariably in one of the known antigenic sites. However, other secondary peaks of hydrophilicity did not correlate very well with known antigenic sites, so that hydrophilicity alone was not sufficiently predictive. Nevertheless, they were able to predict an antigenic site of hepatitis B surface antigen and to verify its existence experimentally.

Fraga (*30*) extended this approach by considering semiempirical "recognition factors," which would reflect the likelihood that hydrophilic amino acids might interact with one another and thus be able to be buried in the protein despite their hydrophilicity. Combining these factors with the Hopp-Woods procedure produced a reasonable correlation between predicted antigenic regions and sites reported to bind antibodies, at least for the limited database of sites considered for each protein. However, in most of these studies of intrinsic factors, the correlation between antigenic sites and the intrinsic property in question is qualitative, not quantitative. It would be useful to define overlaps and statistical significance more quantitatively (*10*).

The advantage of hydrophilicity as a predictor of antigenic sites is that it requires only primary sequence information, not x-ray diffraction crystal structures. However, if the success of the hydrophilicity method is due solely to its correlation with surface exposure, then it is the surface exposure, rather than the hydrophilicity of the residues involved, that is important for antigenicity. Disproportionate increases in antibody binding have been observed during sequential peptide synthesis after addition of lysine residues (*24*), consistent with a major role of these hydrophilic residues in binding. However, such charged residues also contribute greatly to nonspecific binding (*24*). As further support for the importance of charge in antigen-antibody interactions, the classic study of Sela and Mozes and co-workers (*13, 31*) showed that antibodies specific for basic antigens tended to be acidic, whereas those specific for acidic antigens tended to be more basic. As most of the antibody charge variability is in the hypervariable segments of the variable region, which contribute to the combining site, this result suggests that antigen-antibody charge complementarity is frequently important for binding. On the other hand, a significant fraction of surface residues can be nonpolar (*28*), and several groups have described the importance of hydrophobic and especially aromatic residues in antigenicity (*13, 24, 32*). For several nonantibody protein-protein interactions, it has been estimated that most of the bonding energy derives from hydrophobic interactions (that is, exclusion of water over a large surface area of contact, with resultant increase in solvent entropy) (*33*). However, as this contribution to the energy does not contribute much to specificity, the specificity of interaction between two proteins (the selectivity with which binding occurs) depends more on the complementarity of the surfaces for hydrogen bonding and van der Waals contacts (*33*), and on more polar interactions.

In conclusion, although hydrophilicity, as a predictor of surface exposure, may have some value in predicting potential antigenic sites on proteins, it is necessary to be cautious about inferring any mechanistic implication for the biophysics of antigen-antibody interactions.

Mobility. Several groups have proposed that segmental mobility of the polypeptide backbone in portions of protein molecules may contribute to the antigenicity of these sites (*34, 35*). In contrast to hydrophilicity, this parameter may be more important for its theoretical implications in protein dynamics and protein-protein interactions than for its practical predictive value. The studies so far have depended on temperature factors or B values from x-ray diffraction data, which indicate the degree of atomic motion within the crystal structure. Such values can be obtained only from highly refined x-ray crystallographic data, available for only a handful of proteins. Although useful information may also be obtainable from temperature-dependent nuclear magnetic resonance studies (*36*), these, too, are difficult and not readily available for most proteins. Certainly, the method is not applicable to cases in which only the primary sequence is known, although exon-intron boundaries may be indicative of at least some mobile regions (*35*).

Second, for methodological reasons, mobility has been examined only for segmental antigenic sites. Therefore, the large fraction of sites that are of the assembled topographic type are excluded from consideration and would be very difficult to analyze or predict by this approach.

The issue of mobility has been studied from two very different approaches, the

binding of antibodies against peptides and that of antibodies against proteins. These must be considered separately. First, Lerner and colleagues (*1, 37*) were surprised by the high frequency with which antibodies to synthetic peptides too short to have much native structure could nevertheless bind to the intact (although not always fully native) protein. If only a small fraction of the peptide molecules used for immunization were folded in a native-like conformation at any given time, a much smaller proportion of the antibodies might be expected to cross-react with the native protein. Yet, of 38 monoclonal antibodies to six peptides from four proteins, 24 reacted with the intact protein either in solution (where it should be native) or on nitrocellulose sheets (where it would likely not be native, but would also not be as random in conformation as the peptide) (*37*). One explanation considered (*37*) was that the protein structure was flexible enough to fit itself into an antibody combining site specific for a more unfolded conformation of the immunizing peptide (an induced fit hypothesis) (*38*) (see Fig. 2). Tainer *et al.* (*35*), using antisera against 12 overlapping peptides (of about ten residues each) from myohemerythrin, found an extremely good correlation between mobility of a segment of the native protein and the ability of antibodies to the corresponding peptide to bind the native protein in solution. Of course, surface segments of polypeptide may be more mobile than internal ones, but the correlation was not just with surface accessibility. Two regions of low mobility were as exposed as several of the more mobile regions, and yet bound much less to the antisera to the corresponding peptide. It was concluded that atomic mobility of a segment of the native protein was a critical factor in determining the ability of antibodies against short peptides corresponding to that segment to bind to the native protein.

This conclusion is a statement about the specificity and cross-reactivity of antibodies to peptides, not a statement about the immunogenicity of sites on the native protein. As such, it is compatible with much other data (*14*). This approach may be useful in the selection of peptides to elicit antibodies cross-reactive with the native protein, but it does not predict which sites of the native molecule will elicit antibodies when the native protein is used as an immunogen. However, for the purpose of making site-specific antibodies as probes of the native molecule or for production of synthetic vaccines, it is sufficient to be able to select useful

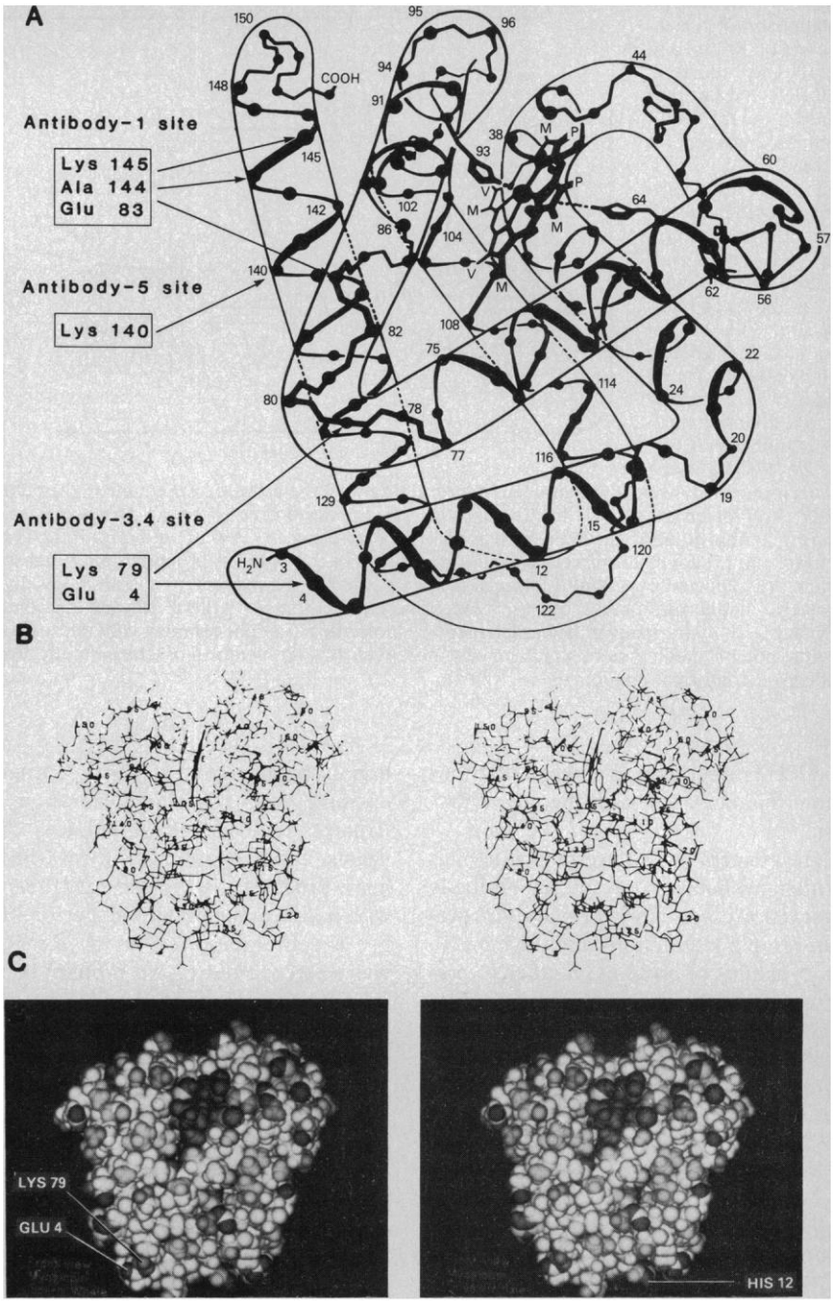

Fig. 1. Three ways of viewing the same face of sperm whale myoglobin, illustrating antigenic sites. (A) A line drawing representing the three-dimensional structure of the α-carbon backbone, modified from Dickerson (*73*). The numbers indicate the amino acid residue positions. M, V, and P are the methyl, vinyl, and propionyl groups of the heme. Side chains are omitted except for two histidines (64 and 93) involved with the heme. The residues involved in binding three conformation-specific monoclonal antibodies (*11*) are shown. Sites of antibodies 1 and 3.4 are assembled topographic sites. Although only one residue in the site has been defined, the binding of antibody 5 is also highly conformation-dependent. Other monoclonal antibodies, to beef myoglobin, bind assembled topographic sites that include residues 74, 87, and 142 and residues 34, 53, and 113, as described by East *et al.* (*11*). Five regions of segmental sites include residues 15 to 29, 56 to 69, 70 to 76, 139 to 146, and 147 to 153 as described by Crumpton and Wilkinson (*16*) or residues 15 to 22, 56 to 62, 94 to 99, 113 to 119, and 145 to 151 as described by Atassi (*4*). Longer peptides outside the latter five regions, which bind a significant fraction of antibodies (*24*), include residues 25 to 55 and 72 to 88 of beef myoglobin. (B) Computer-generated stereo view of the same face of sperm whale myoglobin, showing the amino acid side chains in place (*74*) based on the x-ray crystallographic coordinates of Takano (*75*). (C) Computer-generated stereo space-filling model of sperm whale myoglobin (*74*) and modified from Berzofsky *et al.* (*11*). Stereopairs may be viewed in three dimensions with a stereo viewer. The carboxyl oxygens are shaded darkest, followed by the heme and aromatic carbons, aliphatic side-chain carbons, noncarboxylic oxygens, primary amino groups, and other nitrogens. The backbone and side chains of nonaliphatic residues, except for functional groups, are shown in white. Residues 4, 12, and 79 in the assembled topographic site of antibody 3.4 are indicated. As one progresses from (A) through (B) to (C), it becomes progressively harder to discern the backbone helices, and what is seen is a surface similar to that seen by antibody.

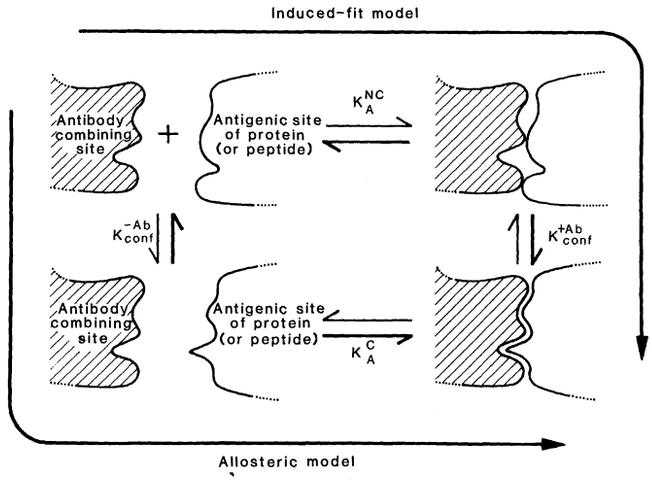

Fig. 2. Schematic comparison of the induced fit (*38*) (top and right) and allosteric (*40*) (left and bottom) models of the equilibriums between an antibody and a protein or peptide antigen whose predominant conformation (top left) does not make a best fit with the antibody combining site. The antigen can achieve greater complementarity by undergoing a conformation change (vertical equilibriums) (*14, 15, 38, 40*). K_A is the association constant for the bimolecular interaction with the antigen in the noncomplementary (NC) or complementary (C) forms. K_{conf} is the equilibrium constant for the unimolecular conformational equilibrium in the absence (−Ab) or presence (+Ab) of antibody. Bold arrows indicate the predominant direction of the reactions. In the induced fit model, an initial weak binding is stabilized by a conformational change in the antigen induced by the antibody. In the allosteric model, an unfavorable conformational equilibrium is "pulled" toward the complementary form by trapping of the complementary molecules in a tight complex with the antibody. A more complicated version, which may be more realistic (*45*), would allow the antibody also to undergo a conformational change.

peptides, rather than to make predictions about immunization with the native protein.

The concept is also useful as it applies to the mechanisms of antigen-antibody interaction and, more broadly, to protein-protein interactions and the molecular dynamics of proteins. Studies of proteins in solution, such as by nuclear magnetic resonance, paint a picture of proteins that is much more mobile and fluctuating than that envisioned from the structure in crystals (*36, 39*). This mobility enables an antibody specific for an unfolded peptide conformation to bind to the intact protein, but at an energy cost. Studies measuring the affinity of antipeptide antibodies for binding to the native protein have found it to be lower than the affinity for the peptide itself by two to three orders of magnitude (*14*). This energy may be required to induce the optimal conformation complementary to the antibody combining site, as in the induced fit model (Fig. 2). Alternatively, it may be the entropic energy of locking in one conformation out of the many that are in equilibrium in solution before the antibody binds. This may be viewed as a type of allosteric model (*40*) (Fig. 2). Either way, since these mechanisms are two paths between the same end points (Fig. 2), the net change in free energy must be the same by the laws of thermodynamics.

Thus, antibodies may be useful for determining conformation (*41*) and dynamic mobility of proteins (*12*) and also for inducing (or stabilizing) conformational changes in a protein. A graphic example was the demonstration by Crumpton (*42*) that antibodies made against apomyoglobin (the form with the heme prosthetic group removed) reacted with native myoglobin to induce or stabilize a conformation in which the heme was squeezed out of the protein; brown myoglobin, precipitated with the antibodies, formed a white precipitate. More recently, monoclonal antibodies to native myoglobin have been found to induce or stabilize a conformation that favors the low-spin over the high-spin electronic state of the ferric heme iron, as measured by optical spectroscopy (*12*). Also, antibodies that bind the oxyhemoglobin conformation with higher affinity than the deoxy conformation enhance the oxygen affinity of human hemoglobin (*43*). Thus, antibodies may act as allosteric effectors and so alter the function of their target antigens.

A second approach to the question of mobility of antigenic structures was to examine the mobility of sites bound by antibodies raised against the native protein. Westhof *et al.* (*34*) compared the atomic mobility of segments of the polypeptide backbone of tobacco mosaic virus protein (as determined from x-ray crystallography) with the known antigenic sites (as determined from antibodies raised against the native protein and mapped to particular segments by binding to synthetic peptides). Only segmental sites could be examined by this approach. Nevertheless, in a plot of mobility (temperature factor) versus sequence, seven of seven sites corresponded to local maxima of mobility, and six of the seven corresponded to major peaks. Moreover, the correlation with mobility seemed better than with surface accessibility, as three hexapeptides that did not bind antibodies corresponded to regions that were exposed but not more than average in mobility. However, subsequently longer peptides corresponding to regions of low mobility were found to bind to antibodies from several antisera against the native protein (*25*). The longer peptides span a broad stretch of low mobility from about residues 115 to 150, to which antibodies had not been detected with the hexapeptides. Thus, the exposed regions of low mobility in the native protein were immunogenic, but the antibodies specific for these sites did not cross-react with the short hexapeptides. Similarly, a good correlation was found between reported segmental antigenic sites for myoglobin and lysozyme and the segmental mobility of these proteins (*34*). However, for myoglobin, seven of the nine sites examined were based on short peptides six to eight residues long (*4*). Thus, with this limited database, antibodies cross-reactive to short peptides were the major category of antibodies studied. Combining the two studies, these investigators have concluded that the entire surface of the tobacco mosaic virus protein is antigenic, and suggest that no single criterion can be used to distinguish nonantigenic from antigenic regions (*25*). As noted earlier, the sites found depend considerably on the probes used in the investigation. Mobility in this case, as for the antipeptide antibodies, is important in determining cross-reactivity rather than immunogenicity. Antibodies to more mobile regions of native proteins are more likely to cross-react with short peptides than are antibodies to less mobile regions, but both are produced on immunization with the native protein. Indeed, if native molecules bearing mutational changes are used as probes, instead of fragments, another category of antibodies is frequently found. These antibodies react with assembled topographic sites that depend strongly on the maintenance of the local tertiary structure (*5*). Thus, we cannot conclude that mobile segments are inherently more immunogenic or antigenic in the native protein than less mobile segments, although it is possible that the relative frequency of antibodies to different regions is affected by their mobility.

Nevertheless, the studies of mobility of antigenic sites provide important insights. In order to use peptides on col-

umns to fractionate antisera against the native protein, it is important to choose peptides corresponding to more mobile regions or to use longer peptides, and to know which antibodies may be missed in such a procedure. Secondly, as discussed by Westhof *et al.* (*34*), antigenic flexibility may contribute to the ability of the immune system to provide a defense against such an enormous diversity of antigens. The clonal selection theory (*44*), for which an enormous body of evidence has been amassed, states that the specificity of antibodies on B lymphocytes is determined genetically before antigen enters the system. The antigen activates and expands those clones of B cells that bear antibodies capable of binding the antigen. However, it is unlikely that there will be preformed antibodies on B cells that can bind every possible conformation of every polypeptide sequence. If the protein antigens are flexible, they are more likely to find some antibody for which they can achieve a satisfactory (induced) fit (*34, 35*). This possibility expands the potential repertoire of antibodies, but at the cost of reduced affinity.

Finally, the implications of the mobility hypothesis for the biophysics of antigen-antibody and, in general, protein-protein interactions may be of great importance. Neither the antigen nor the antibody combining site can be viewed as a static structure. X-ray diffraction of a series of fluorescent ligands interacting with the combining site of an immunoglobulin light-chain dimer showed that both side chains and polypeptide backbone of the antibody combining site move to accommodate different ligands (*45*). Thus, structural flexibility of both the protein antigen and the antibody may be valuable to achieve optimum complementarity.

Factors Extrinsic to Antigen Influencing Immunogenicity of Specific Sites

Although intrinsic factors may determine the repertoire of potential antigenic sites in a protein antigen, only a subset of these sites will elicit antibodies when any given person or animal is immunized. For example, two peptides reported to bind antibodies made against sperm whale myoglobin in rabbits, goats, and mice (*4*) did not bind any detectable antimyoglobin antibodies made in two other goats, a sheep, and several high-responder strains of mice (*46*). Thus, factors in the host being immunized, as distinct from any structural features inherent in the antigen, may be of paramount importance in determining the outcome of any immunization. The discussion here will be limited to a summary of the impact of these factors on antibody specificity.

Tolerance to self. Self-tolerance is one of the fundamental properties of the immune system (*47*). The ability to break tolerance by experimental manipulation and the appearance of autoantibodies in certain pathologic states indicates that the potential structural gene repertoire is present in the genome to make antibodies that will react with self. Nevertheless, under normal circumstances, antibodies to host proteins are not made. Therefore, when the immunogen is a mammalian or avian protein homologous to a protein of the host, antibodies will be made primarily to those sites that differ from those of the host protein (*5*). For example, when rabbits were immunized with guanaco, mouse, or horse cytochrome c, the subpopulations of antibodies detected could mostly be accounted for by binding to the sites at which rabbit cytochrome c differs from the immunogen (*48*).

A second example is that of antibodies to beef myoglobin raised in various species (*49*). The antibodies raised in rabbits, dogs, and chickens bound almost equally well to bovine or sheep myoglobin, whereas the antibodies raised in sheep distinguished strongly between beef and sheep myoglobin, which differ at only 6 of 153 residues. Thus, the sheep's self-tolerance led to antibodies specific for sites at which beef and sheep myoglobin differ, whereas most of the antibodies raised in the other species bound sites shared by beef and sheep myoglobins. Differences between the immunogen and homologous host proteins are more important than anything inherent in the structure of the immunogen in determining the outcome of an immune response, at least for antigens which have homologues in the host. The same should apply to monoclonal antibodies as well, although an occasional autoantibody may be isolated.

Immune response (Ir) *genes.* Ir genes are genes that regulate the ability of an individual to make an immune response to a specific antigen (*6*). Antigen specificity is a key part of the definition, as the gene responsible for a broad-spectrum immunodeficiency disease would not be called an Ir gene. Most of the immunoregulatory genes that have proved to be antigen-specific, and thus qualify as Ir genes, are part of the major complex of genes encoding transplantation antigens, known as the major histocompatibility complex (MHC). Structural mutations have definitively shown that the Ir genes are actually the structural genes for MHC antigens (*6*). Their mechanism of action appears to be involved with the way in which T lymphocytes are activated, not by free antigen in solution, but only by a combination of the antigen with an appropriate MHC antigen on the surface of another cell. The Ir genes thus determine which T lymphocytes are activated, although it is still debated whether they accomplish this through a direct, specific interaction with the antigen or through an effect on the available repertoire of T cells.

In the case of antibody responses, Ir genes act indirectly, by influencing the level and specificity of helper T lymphocytes, which are required for B cells to be activated to make antibodies to protein antigens. Nevertheless, probably via this indirect route, Ir genes influence the site specificity of antibodies raised to protein antigens (for example, see Fig. 3). Mice of strains B10 and B10.A are genetically identical except for their MHC genes. Although B10 is initially a lower responder, after three immunizations with staphylococcal nuclease both strains make equivalent levels of antinuclease antibodies (*50*). However, only the B10.A sera contain antibodies that bind fragment 99–149; the B10 sera do not detectably bind this fragment (*50*). A second example is the case of antibody responses against sperm whale myoglobin in three strains of mice, B10.D2, B10.A, and B10.BR, which are again congenic (identical except for their MHC genes). B10.D2 mice are high responders to sperm whale myoglobin and make antibodies to a number of sites, including antibodies which bind to fragment 132–153 (*51*). B10.BR mice are low responders and make little total antibody, of which almost none binds to fragment 132–153. The B10.A mice are intermediate in total magnitude of response, but the antibodies to myoglobin that are made contain no more antibodies binding to fragment 132–153 than do those of the low responder B10.BR mice (*51*) (Fig. 3). Thus, Ir genes control the site specificity of antibodies produced, not just the total magnitude of the response. These congenic strains of mice do not differ in their myoglobin structural genes, and they have no protein homologous to staphylococcal nuclease. Therefore, the differences in antibodies produced are not due to tolerance to homologous host proteins. The first case of this type described was for antibodies to a synthetic random polymer of three amino acids (*52*). Thus, Ir genes are a second host factor that strongly influences which

sites on a protein are immunogenic in a given animal. In an outbred species such as man, every individual will have a different complement of Ir genes. Other examples of potential clinical relevance include Ir gene control of the specificity of antibodies to hepatitis B surface antigen (*53*) and to peptides of myelin basic protein (*54*). In this context, the frequency of low responders to a peptide vaccine bearing a single antigenic site would be expected to be greater than the frequency of low responders to a whole protein with many sites, each under different Ir gene control (*51*). Perhaps mixtures of peptides could circumvent this potential problem.

Helper and suppressor T-lymphocyte specificity and T-B reciprocity. B lymphocytes specific for protein antigens generally require "help" from a T lymphocyte specific for the same antigen molecule (although not necessarily the same site) in order to be activated to secrete antibodies. Ir genes act on the magnitude and specificity of T-cell help, but to explain their action on antibody specificity, it is necessary to postulate that the specificity of T cells elicited influences the specificity of antibodies produced (*55*). Because Ir gene products that influence T-cell specificity are expressed on B lymphocytes, and the resulting T-cell specificity appears to lead to selective activation of a subset of B lymphocytes, this postulated reciprocal effect of T and B lymphocytes on each other's specificity has been termed "T-B reciprocity" (*55*). Evidence supporting this notion is still largely indirect but has been found for a number of antigens (*6, 50, 51, 56*).

There are also several cases in which Ir genetic low responsiveness to an antigen is accompanied by active suppression of the response by suppressor T lymphocytes (*6, 8, 57*). Such suppressor T cells have been found, in at least one case, to act selectively on a subset of helper T cells specific for certain antigenic sites (*58*). If suppressor cells influence the specificity of the helper population and this, in turn, selects the B cell repertoire to be activated, suppressor T cells might influence antibody specificity as well. However, a direct demonstration is lacking.

The effect of T-cell helper specificity on the specificity of antibodies produced need not be limited to Ir genes. Anything else which affects T-cell specificity could influence the antibody response. One example is antigen processing. For T lymphocytes to be activated by antigen "presented" on the surface of another cell rather than free in solution, the antigen must first be proteolytically digested or processed by the presenting cell (*59*). The purpose of this processing may be to unfold the protein and expose sites that are necessary for interaction with the MHC antigen or the plasma membrane of the presenting cell (*60*). Amino acid substitutions that affect processing can affect the T-cell response. For instance, Shastri *et al.* (*61*) studied mouse T-cell clones that responded to a tryptic fragment of residues 74 to 96 from either chicken or ring-necked pheasant lysozyme, which share this part of their sequence. Yet, when native lysozymes were used, only that from pheasant would stimulate the clones. Cleaving a small segment from both the amino and carboxyl termini of the lysozymes made the two lysozymes equivalent in stimulatory activity. Thus, a difference far from the antigenic site seemed to influence the proteolytic processing of the chicken lysozyme so as to make the antigenic site seen by the T-cell clone unavailable. If differences in antigen processing lead to differences in helper T-cell specificity, they may also affect antibody specificity.

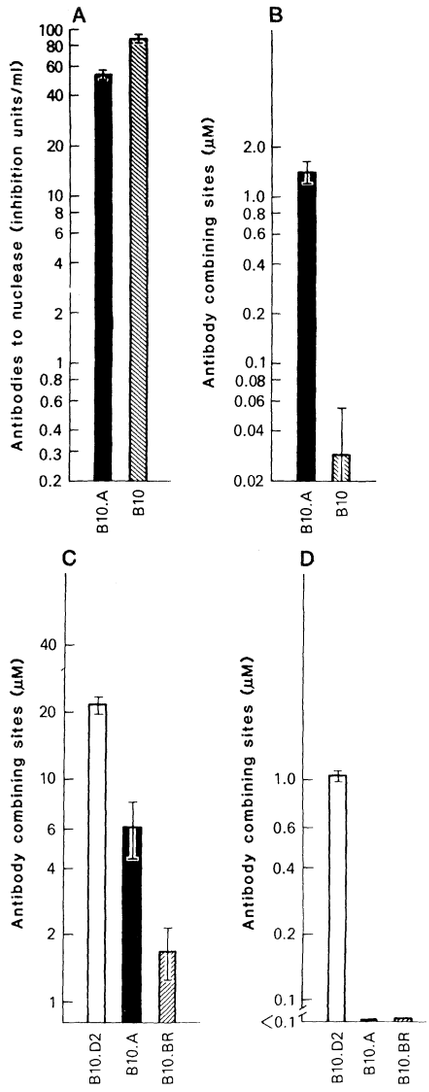

Fig. 3. Effect of Ir genes on antibody specificity (*50, 51*). (A and B) Antibodies made by immunizing three times with native staphylococcal nuclease and testing the same sera for binding to either native nuclease (A) or fragment 99–149 of nuclease (B). The congenic strains of mice, B10.A and B10, differ only in their major histocompatibility genes, including Ir genes. Although the total level of antinuclease antibodies is comparable, only the B10.A sera contain antibodies binding fragment 99–149. (C and D) Antibodies made by immunizing with native sperm whale myoglobin and testing the same sera for binding to native myoglobin (C) or the cyanogen-bromide cleavage fragment 132–153 (D) by radiobinding assay. Means and geometric standard error of the means are shown.

The requirements for helper T-cell recognition also have important implications for peptide vaccine development. In order to ensure an anamnestic response when the recipient is challenged with whole antigen or virus, the peptide must contain not only a site that can induce antibody but also a site that will elicit helper T cells. This point has recently been demonstrated experimentally for peptide antigens from tobacco mosaic virus protein (*63*).

Idiotype networks. The idiotype of an antibody, first described in 1963 (*64*), is that part which is antigenically unique, as determined by a second set of antibodies made against the first one. Idiotypes, therefore, are in the variable region, mostly in or near the combining site. As the same combining site structure may determine both the antigen

specificity and the idiotype of an antibody, there should be some correlation between specificity and idiotype, although not a perfect one. Jerne (*7*) proposed that a system of antibodies and other antibodies against them (anti-idiotypes) could act as a self-contained homeostatic network. The role of antigen was to perturb this internal network. Many studies have supported the existence of such networks (*7*). These networks may involve idiotype-specific helper T cells as well as antibodies. If networks regulate idiotype, they must influence the prevalence of different combining sites and thus antibody specificity. This regulatory mechanism may even be subject to nongenetic inheritance. Rubinstein *et al.* (*65*) injected newborn mice with A48, a monoclonal antibody to levan with a rare idiotype not normally found in antilevan antisera. When these mice were immunized with bacterial levan 4 weeks later, the antibody response was dominated by the normally rare idiotype of A48. The predominance of this rare idiotype could be transferred to other mice with helper T cells from the treated mice. Thus, the prevalence at birth of this idiotype induced complementary idiotype-specific helper cells which altered the spectrum of idiotypes produced on immunization later in life. Since immunoglobulin G crosses the placenta, Rubinstein *et al.* (*65*) suggested that the idiotype history of the mother could be passed on to her offspring, a form of nongenetic inheritance that may have some protective advantage.

The idiotype expressed in a given immune response can also be influenced by MHC-linked Ir genes (*66*). Thus, there is an interface at which these two major regulatory mechanisms can interact. The net effect is that antigen specificity of the antibody response depends on both the genetics and the immune history of the host, and even on the immune history of the previous generation.

Structural gene repertoire. The diversity of the potential antibody repertoire depends on the combinatorial joining of structural genes for several elements which make up the antibody combining site (V_H, D, and J_H on the heavy chain and V_L and J_L on the light chain). There is a further diversity generated by combinatorial pairing of heavy and light chains and by somatic mutation (*67*). Loh *et al.* (*68*) showed that the V_H that dominates the antibody response to the 4-hydroxy-3-nitrophenyl acetyl group (NP) in one strain of mice (BALB/c) comes from a completely different family of V_H genes from that which dominates the response to NP in another strain of mice, C57BL/6. Thus, there appears to have been enough genetic divergence in V_H between mouse strains to influence the specificity of antibody responses to the same small antigen.

Stochastic effects and clonal preemption. The antibody response to *p*-azobenzene arsonate (ABA) in A/J mice is dominated by a major idiotype (*69*). However, the repertoire of B-cell clones available prior to immunization is quite diverse, and the dominance of a particular idiotype in this response is due to selective forces during the immune response itself, such as affinity of B cells for antigen (*70*). Later in the immune response, clones lacking the dominant idiotype may arise by somatic mutation with affinities higher than that of the dominant idiotype. However, it is supposed that once the clones with dominant idiotype have gained a head start at expansion, they preempt the response, perhaps by consuming antigen, so that the other clones cannot catch up. An earlier study of the same ABA response supports the idea that clones with a head start can preempt the response (*71*). When ABA-specific B cells expressing idiotypes other than the dominant one were transferred into naive A/J mice prior to immunization, these clones now dominated the ABA response, and the major idiotype that dominates in unmanipulated A/J mice was suppressed (*71*). Thus, in addition to all the regulatory mechanisms that influence antibody specificity, there is a stochastic element as well that can result in animal to animal variability. In each animal, the clone that gets the antigen first may preempt the response. However, there are cases, such as the response to the hapten 4-hydroxy-3-nitrophenyl acetyl, in which an idiotype that predominates early in the response becomes negligible after repeated immunizations, when other idiotypes take over (*72*). Thus, other regulatory mechanisms can override clonal preemption. Indeed, all of the host regulatory mechanisms act in concert to determine the ultimate magnitude and specificity of the observed antibody response.

Conclusions

In summary, the weight of the evidence is that virtually the entire accessible surface of a protein may be antigenic. Intrinsic features such as hydrophilicity of stretches of amino acids help predict accessibility from primary sequence. Mobility of a polypeptide segment of a protein, in the few cases in which mobility data can be obtained, is useful for predicting which segmental sites of a protein are likely to elicit antibodies that cross-react with the corresponding peptide or are likely to cross-react with antibodies made against a synthetic peptide of like sequence. Flexibility of an antigenic site also has important implications for the dynamics of antigen-antibody and protein-protein interactions in general. However, mobility probably does not enhance immunogenicity per se, and these studies cannot be applied to the large number of antibodies that bind assembled topographic sites—that is, sites comprising amino acid residues far apart in the primary sequence but brought together on the surface by the folding of the protein in its native conformation. Of all the potential antigenic sites of a protein, only a subset will be immunogenic in any individual host. Thus, host factors (extrinsic to the antigen molecule) are critical to immunogenicity of a potential antigenic site. These host regulatory factors include self-tolerance, immune response genes, specificity of helper and suppressor T cells, antigen processing, idiotype networks, the host's structural antibody gene repertoire, and clonal preemption. While these extrinsic factors make it difficult to predict accurately the structure of individual antigenic sites from a knowledge of the protein structure alone, it is still possible to make worthwhile selections of peptide segments likely to be useful as probes or as synthetic vaccines. Indeed, one of the powers of the synthetic immunogen approach is that peptides may be immunogenic in a given host even when the corresponding site in the native protein is not (*1, 8, 61*). Thus, synthetic peptide immunogens may circumvent some extrinsic factors.

References and Notes

1. R. A. Lerner, *Nature (London)* **299**, 592 (1982); *Adv. Immunol.* **36**, 1 (1984).
2. R. Arnon, in *Impact of Protein Chemistry on the Biomedical Sciences*, A. N. Schechter, A. Dean, R. F. Goldberger, Eds. (Academic Press, Orlando, Fla., 1984), p. 187.
3. K. Landsteiner, *The Specificity of Serological Reactions* (Thomas, Springfield, Ill., 1936).
4. M. Z. Atassi, *Immunochemistry* **12**, 423 (1975); S. S. Twining, H. Lehmann, M. Z. Atassi, *Biochem. J.* **191**, 681 (1980); S. S. Twining, C. S. David, M. Z. Atassi, *Mol. Immunol.* **18**, 447 (1981).
5. D. C. Benjamin *et al.*, *Annu. Rev. Immunol.* **2**, 67 (1984).
6. B. Benacerraf and R. N. Germain, *Immunol. Rev.* **38**, 71 (1978); A. S. Rosenthal, *ibid.* **40**, 136 (1978); J. A. Berzofsky, in *Biological Regulation and Development*, R. F. Goldberger, Ed. (Plenum, New York, 1980), vol. 2, pp. 467–594; R. H. Schwartz, in *Ia Antigens*, vol. 1, *Mice*, S. Ferrone and C. S. David, Eds. (CRC Press, Boca Raton, Fla., 1982), pp. 161–218; J. A. Berzofsky, in *The Antigens*, M. Sela, Ed. (Academic Press, New York, in press).
7. N. K. Jerne, *Ann. Immunol. (Paris)* **125C**, 373 (1974); C. A. Bona and B. Pernis, in *Fundamen-*

tal Immunology, W. E. Paul, Ed. (Raven, New York, 1984), pp. 577–592.
8. E. E. Sercarz *et al.*, *Immunol. Rev.* **39**, 108 (1978); J. W. Goodman and E. E. Sercarz, *Annu. Rev. Immunol.* **1**, 465 (1983).
9. C. G. Fathman and J. G. Frelinger, *Annu. Rev. Immunol.* **1**, 633 (1983).
10. C. DeLisi and J. A. Berzofsky, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
11. J. A. Berzofsky *et al.*, *J. Biol. Chem.* **257**, 3189 (1982); I. J. East *et al.*, *ibid.*, p. 3199.
12. J. A. Berzofsky, in *Monoclonal and Antiidiotypic Antibodies: Probes for Receptor Structure and Function*, J. C. Venter, C. M. Fraser, J. Lindstrom, Eds. (Liss, New York, 1984), pp. 1–19.
13. M. Sela, *Science* **166**, 1365 (1969).
14. D. H. Sachs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3790 (1972); B. Furie *et al.*, *J. Mol. Biol.* **92**, 497 (1975); J. G. R. Hurrell, J. A. Smith, S. J. Leach, *Biochemistry* **16**, 175 (1977); M. J. Darsley and A. R. Rees, *EMBO J.* **4**, 383 (1985); A. Hirayama, Y. Takagaki, F. Karush, *J. Immunol.* **134**, 3241 (1985); I. A. Wilson, personal communication.
15. J. A. Berzofsky and A. N. Schechter, *Mol. Immunol.* **18**, 751 (1981).
16. M. J. Crumpton and J. M. Wilkinson, *Biochem. J.* **94**, 545 (1965).
17. G. Köhler and C. Milstein, *Nature (London)* **256**, 495 (1975); F. Melchers, M. Potter, N. L. Warner, *Lymphocyte Hybridomas* (Springer-Verlag, Berlin, 1978); R. H. Kennett, T. J. McKearn, K. B. Bechtol, *Monoclonal Antibodies. Hybridomas: A New Dimension in Biological Analysis* (Plenum, New York, 1980).
18. G. Lando, J. A. Berzofsky, M. Reichlin, *J. Immunol.* **129**, 206 (1982).
19. J. A. Berzofsky and I. J. Berkower, in *Fundamental Immunology*, W. E. Paul, Ed. (Raven, New York, 1984), pp. 595–644.
20. S. J. Smith-Gill *et al.*, *J. Immunol.* **128**, 314 (1982); S. J. Smith-Gill, T. B. Lavoie, C. R. Mainhart, *ibid.* **133**, 384 (1984); D. W. Metzger *et al.*, *Eur. J. Immunol.* **14**, 87 (1984); A. G. Amit, R. A. Mariuzza, S. E. V. Phillips, R. J. Poljak, *Nature (London)* **313**, 156 (1985); M. Z. Atassi, *Immunochemistry* **15**, 909 (1978); _____ and C.-L. Lee, *Biochem. J.* **171**, 429 (1978).
21. The validity of defining sites by comparing evolutionary sequence variants was recently confirmed by P. V. Hornbeck and A. C. Wilson [*Biochemistry* **23**, 998 (1984)], who studied the affinity of lysozymes for the substrate analog chitotriose. Only amino acid substitutions within the binding site affected affinity, whereas even multiple substitutions outside the site had no significant effect.
22. D. C. Wiley, I. A. Wilson, J. J. Skehel, *Nature (London)* **289**, 373 (1981); P. A. Underwood, *J. Gen. Virol.* **62**, 153 (1982); L. M. Staudt and W. Gerhard, *J. Exp. Med.* **157**, 687 (1983); P. M. Colman, J. N. Varghese, W. G. Laver, *Nature (London)* **303**, 41 (1983).
23. Z. Al Moudallal, J. P. Briand, M. H. V. Van Regenmortel, *EMBO J.* **1**, 1005 (1982); D. Altschuh *et al.*, *Mol. Immunol.* **22**, 329 (1985).
24. S. J. Leach, *Biopolymers* **22**, 425 (1983); P. E. E. Todd, I. J. East, S. J. Leach, *Trends Biochem. Sci.* **7**, 212 (1982); P.-T. Shi *et al.*, *Mol. Immunol.* **21**, 489 (1984).
25. Z. Al Moudallal, J. P. Briand, M. H. V. Van Regenmortel, *EMBO J.*, in press.
26. T. J. White, I. M. Ibrahimi, A. C. Wilson, *Nature (London)* **274**, 92 (1978).
27. D. C. Benjamin and J. M. Teale, *J. Biol. Chem.* **253**, 8087 (1978); D. C. Benjamin, L. A. Daigle, R. L. Riley, in *Protein Conformation as an Immunological Signal*, F. Celada *et al.*, Eds. (Plenum, New York, 1983), pp. 261–280; T. Peters, Jr., R. C. Feldhoff, R. G. Reed, *J. Biol. Chem.* **252**, 8464 (1977); T. Kamiyama, *Immunochemistry* **14**, 91 (1977); N. Doyen, A. J. Pesce, C. Lapresle, *J. Biol. Chem.* **257**, 2770 (1982).
28. Accessibility of atoms of a protein to solvent or solute molecules of various sizes can be defined quantitatively [B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971); M. L. Connolly, *Science* **221**, 709 (1983)].
29. T. P. Hopp and K. R. Woods, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3824 (1981); *Mol. Immunol.* **20**, 483 (1983).
30. S. Fraga, *Can. J. Chem.* **60**, 2606 (1982).
31. M. Sela and E. Mozes, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 445 (1966).
32. E. Benjamini *et al.*, *Biochemistry* **7**, 1261 (1968).
33. C. Chothia and J. Janin, *Nature (London)* **256**, 705 (1975).
34. E. Westhof *et al.*, *ibid.* **311**, 123 (1984).
35. J. A. Tainer *et al.*, *ibid.* **312**, 127 (1984); J. A. Tainer *et al.*, *Annu. Rev. Immunol.* **3**, 501 (1985).
36. G. R. Moore and R. J. P. Williams, *Eur. J. Biochem.* **103**, 543 (1980).
37. H. L. Niman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4949 (1983).
38. D. E. Koshland, Jr., G. Nemethy, D. Filmer, *Biochemistry* **5**, 365 (1966).
39. F. R. N. Gurd and R. M. Rothgeb, *Adv. Protein Chem.* **33**, 73 (1979); M. Levitt, *J. Mol. Biol.* **168**, 621 (1983).
40. J. Monod, J. Wyman, J.-P. Changeux, *J. Mol. Biol.* **12**, 88 (1965).
41. R. M. Lewis, B. C. Furie, B. Furie, *Biochemistry* **22**, 948 (1983); J. Owens *et al.*, *J. Biol. Chem.* **259**, 13800 (1984).
42. M. J. Crumpton, in *The Antigens*, M. Sela, Ed. (Academic Press, New York, 1974), vol. 2, pp. 1–78.
43. J. Dean and A. N. Schechter, *J. Biol. Chem.* **254**, 9185 (1979).
44. F. M. Burnet, *The Clonal Selection Theory of Immunity* (Vanderbilt Univ. Press, Nashville, 1959).
45. A. B. Edmundson, K. R. Ely, J. N. Herron, *Mol. Immunol.* **21**, 561 (1984).
46. J. A. Berzofsky *et al.*, in *Protein Conformation as Immunological Signal*, F. Celada *et al.*, Eds. (Plenum, New York, 1983), pp. 165–180.
47. G. J. V. Nossal, *Annu. Rev. Immunol.* **1**, 33 (1983).
48. G. J. Urbanski and E. Margoliash, *J. Immunol.* **118**, 1170 (1977); R. Jemmerson and E. Margoliash, *J. Biol. Chem.* **254**, 12706 (1979).
49. H. M. Cooper *et al.*, *Mol. Immunol.* **21**, 479 (1984).
50. J. A. Berzofsky *et al.*, *J. Exp. Med.* **145**, 111 (1977); J. A. Berzofsky *et al.*, *ibid.*, p. 123; J. A. Berzofsky *et al.*, *Adv. Exp. Med. Biol.* **98**, 241 (1978); J. A. Berzofsky, unpublished observations.
51. J. A. Berzofsky, *J. Immunol.* **120**, 360 (1978); _____, L. K. Richman, D. J. Killion, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4046 (1979); Y. Kohno and J. A. Berzofsky, *J. Immunol.* **128**, 2458 (1982).
52. H. G. Bluestein *et al.*, *J. Exp. Med.* **135**, 98 (1972).
53. D. R. Milich *et al.*, *ibid.* **159**, 41 (1984).
54. R. B. Fritz *et al.*, *J. Immunol.* **134**, 2328 (1985).
55. J. A. Berzofsky, *Surv. Immunol. Res.* **2**, 223 (1983).
56. J. M. Cecka *et al.*, *Eur. J. Immunol.* **6**, 639 (1976); E. Sercarz *et al.*, *Ann. Immunol. (Paris)* **128C**, 599 (1977); A. Campos-Neto, H. Levine, S. F. Schlossman, *Cell. Immunol.* **69**, 128 (1982); F. Celada *et al.*, in *Regulation of the Immune System*, H. Cantor, L. Chess, E. Sercarz, Eds. (Liss, New York, 1984), pp. 637–646; T. A. Ferguson *et al.*, *Cell. Immunol.* **78**, 1 (1983).
57. J. A. Kapp *et al.*, *J. Exp. Med.* **140**, 648 (1974).
58. U. Krzych, A. V. Fowler, E. E. Sercarz, *ibid.*, **162**, 311 (1985).
59. E. R. Unanue, *Annu. Rev. Immunol.* **2**, 395 (1984).
60. J. A. Berzofsky, in *The Year in Immunology 1984–85*, J. M. Cruse and R. E. Lewis, Jr., Eds. (Karger, Basel, 1985), pp. 18–24; H. Z. Streicher *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6831 (1984).
61. N. Shastri, A. Miller, E. E. Sercarz, in *Regulation of the Immune System*, H. Cantor, L. Chess, E. E. Sercarz, Eds. (Liss, New York, 1984), pp. 153–161.
62. K. L. Rock, B. Benacerraf, A. K. Abbas, *J. Exp. Med.* **160**, 1102 (1984); B. A. Malynn and H. H. Wortis, *J. Immunol.* **132**, 2253 (1984).
63. E. Benjamini *et al.*, in *High Technology Route to Virus Vaccines*, G. R. Dreesman, J. G. Bronson, R. Kennedy, Eds. (American Society for Microbiology, Washington, D.C., in press).
64. H. G. Kunkel, M. Mannik, R. C. Williams, *Science* **140**, 1218 (1963); J. Oudin and M. Michel, *C. R. Acad. Sci.* **257**, 805 (1963); K. Rajewsky and T. Takemori, *Annu. Rev. Immunol.* **1**, 569 (1983).
65. L. J. Rubinstein, M. Yeh, C. A. Bona, *J. Exp. Med.* **156**, 506 (1982).
66. H. Kawamura *et al.*, *ibid.* **160**, 659 (1984); U. M. Babu and P. H. Maurer, *ibid.* **154**, 649 (1981); M. C. Bekoff, H. Levine, S. F. Schlossman, *J. Immunol.* **129**, 1173 (1982).
67. S. Tonegawa, *Nature (London)* **302**, 575 (1983).
68. D. Y. Loh *et al.*, *Cell* **33**, 85 (1983).
69. M. F. Gurish and A. Nisonoff, in *Idiotypy in Biology and Medicine*, H. Köhler, J. Urbain, P.-A. Cazenave, Eds. (Academic Press, New York, 1984), pp. 63–88.
70. T. Manser, S.-Y. Huang, M. L. Gefter, *Science* **226**, 1283 (1984); T. Manser *et al.*, *Immunol. Today* **6**, 94 (1985).
71. F. L. Owen and A. Nisonoff, *J. Exp. Med.* **148**, 182 (1978).
72. O. Mäkelä and K. Karjalainen, *Immunol. Rev.* **34**, 119 (1977).
73. R. E. Dickerson, in *The Proteins*, H. Neurath, Ed. (Academic Press, New York, ed. 2, 1964), vol. 2, pp. 603–778.
74. R. J. Feldmann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5409 (1978).
75. T. Takano, *J. Mol. Biol.* **110**, 537 (1977).
76. I am grateful to Drs. David Davies, Charles DeLisi, Arnold Feinstein, Elizabeth Getzoff, Richard A. Lerner, Alan N. Schechter, Eli Sercarz, Sandra Smith-Gill, John Tainer, and Marc Van Regenmortel for critical reading of the manuscript and for very helpful discussions and to R. Feldmann for computer graphics used in Fig. 1.