Cell, Vol. 61, 553–556, May 18, 1990, Copyright © 1990 by Cell Press

Epitopes on Protein Antigens: Misconceptions and Realities

Minireview

W. Graeme Laver,* Gillian M. Air,† Robert G. Webster,‡ and Sandra J. Smith-Gill§
*John Curtin School of Medical Research
Australian National University
Canberra, ACT 2601, Australia
†Department of Microbiology
University of Alabama at Birmingham
Birmingham, Alabama 35294
‡St. Jude Children's Research Hospital
Memphis, Tennessee 38101
§Laboratory of Genetics
National Cancer Institute
Bethesda, Maryland 20892

Of an almost infinite variety of epitopes (or antigenic determinants) on protein molecules, only five complete structures are known. The single method that can determine the complete structure of an epitope is preparation of a complex of a monoclonal antibody Fab fragment with its antigen, crystallization of this complex, and determination of its structure using X-ray diffraction methods. This has now been done for three complexes of Fab–lysozyme (from chicken egg white) and two of Fab–neuraminidase (from influenza virus) (Amit et al., 1986; Sheriff et al., 1987; Padlan et al., 1989 and Tulip et al., 1990). These epitopes all occupy large areas comprised of 15–22 amino acid residues on several surface loops. Antigenicity of these epitopes is absolutely dependent upon conformation of the native proteins.

The term epitope was coined by Niels Jerne in 1960, when he proposed that "an antigen particle carries several epitopes (= surface configurations, single determinants, structural themes, immunogenic elements, haptenic groups, antigenic patterns, specific areas). Hidden epitopes which become immunologically available only after breakage, decomposition, or denaturation of the antigen are cryptotopes." Inherent in this definition is the concept that an epitope occurs on the surface of a native protein, whereas a linear peptide sequence seen after protein unfolding could occur on the inside or the outside of the folded polypeptide. Processed peptides that would be recognized by T cells in association with major histocompatibility complex molecules would actually be included in the latter category and are more appropriately considered cryptotopes rather than T cell epitopes as they are now known. Unfortunately, the term T cell epitope is in common usage to describe the peptide sequence on the original protein, although it is not in this native form that it is recognized by the T cell. In contrast, antibodies do recognize epitopes on native proteins, and the complementary paratope on the antibody is exquisitely specific for the native conformation. We propose that the term epitope be reserved for those structures on native proteins that bind antibodies, and the structures on unfolded proteins and those recognized by T cells should be termed something else.

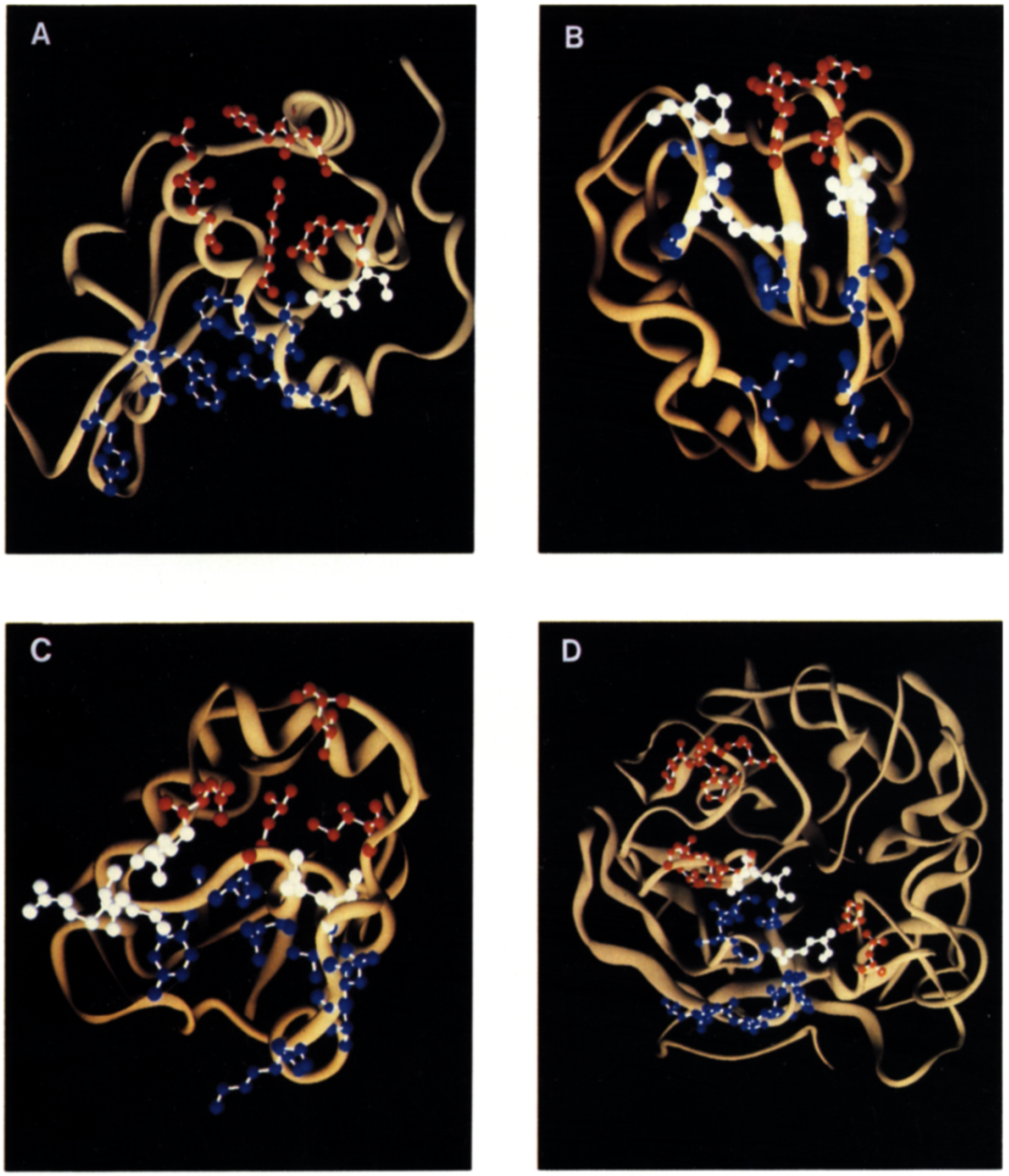
For protein antigens, it was later proposed that epitopes might be subdivided into sequential epitopes (involving a single continuous length of the polypeptide chain) and conformational epitopes, in which several discrete amino acid sequences, widely separated in the primary structure, come together on the surface when the polypeptide chain folds to form the native protein (Sela, 1969). As discussed below, all five of the structurally defined epitopes are of the conformational type. As has been discussed previously (Benjamin et al., 1984), all determinants recognized by antibodies are conformational in that antibodies will bind with measurable affinity only to those molecules presenting the right conformation; "discontinuous" is a more accurate description of nonlinear epitopes since they are assembled from residues from several different portions of the polypeptide chain.

The five epitopes characterized by X-ray crystallography are of the discontinuous type, encompassing two to five surface loops. The highly complementary interface between antigen and antibody is absolutely dependent on the folding of the native protein. The antibody combining site is seen to be an irregular, rather flat surface with protrusions and valleys formed by amino acid side chains. The epitopes each contain between 15 and 22 residues on the antigen in contact with a similar number of residues on the antibody paratope. Although the interpretation of contacts is somewhat arbitrary (and may differ between laboratories), and the contribution of each to the binding energy is unknown, these five examples provide the best descriptions of epitopes on native proteins so far available. The figure shows the visualization of one epitope on influenza virus neuraminidase and three on lysozyme. Each epitope has a buried surface area on the antigen of 650–900 $Å^2$. There are 75–120 hydrogen bonds between the antibody and antigen, as well as salt links and hydrophobic interactions.

In contrast to the consensus that X-ray structures have given of protein epitopes, there still is found in the literature a widely held misconception that epitopes on native proteins consist of segments of about 6 amino acid residues that can be mimicked or mapped utilizing synthetic peptides of a similar length. In fact, numerous studies (e.g., Green et al., 1982) claimed to have localized epitopes on native proteins by studying synthetic peptides corresponding to short linear sequences within the protein. However, the success of such approaches with well-defined antibodies has been limited, and any cross-reactivity seen in these experiments probably represents binding to a proportion of denatured protein (Jemmerson and Blankenfeld, 1989). Similarly, those cases in which peptides appear to give good cross-reactivity with anti-protein antibodies in solution assays probably represent antibodies originally elicited by denatured protein.

The still-prevailing notion that a protein epitope has the size of a hexapeptide came from extrapolation of studies with carbohydrate antigens. A recent article (Goodman, 1989) describes epitopes as spanning 4–7 amino acid residues on the antigen. This concept has also received

Petitioner Merck, Ex. 1184, p. 553
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00003



Ribbon Diagrams Showing the Location of Epitopes

Structures were determined by X-ray diffraction analysis of crystalline complexes of monoclonal antibody Fab fragments and the antigens, hen egg white lysozyme and influenza virus neuraminidase. (A) HyHEL10–lysozyme. (B) HyHEL5–lysozyme. (C) D1.3–lysozyme. (D) NC41–neuraminidase. Residues contacting antibody light chains, heavy chains, or both are shown in red, blue, and white, respectively. Contacts for the Fab–lysozyme complexes are based on Amit et al. (1986), Sheriff et al. (1987), and Padlan et al. (1989), as summarized by Davies et al. (1990). The Fab–neuraminidase contacts are those listed by Tulip et al. (1990). The program used was Ribbons, written by Dr. M. Carson. Models were constructed by Dr. Ming Luo.

apparent support from binding studies with anti-peptide antibodies. These indicate that antibodies generated against peptides would bind to proteins in assays such as ELISA, a test in which there is good opportunity for complete or partial unfolding of the protein antigens (e.g., Green et al., 1982). Perhaps a better term for determinants detected under these conditions is unfoldon. The term epitope should be reserved for those determinants recognized by antibodies on native proteins. How can one determine if a particular antibody binds to an epitope on the native pro-

tein or to an unfoldon? For a protein to have biological activity (e.g., enzymatic activity) it needs to be correctly folded. It follows that if an antibody affects that activity, it must be binding to an epitope on the native protein, i.e., a foldon, and not to a cryptotope or unfoldon on the denatured form. Of course, many antibodies will bind to epitopes on native proteins without affecting biological activity.

On the other hand, immunization with peptides can be very useful for producing antibodies for identification of denatured or unfolded molecules (for example, on a gel or in a bacterial expression library). It must be emphasized that these antibodies are not recognizing epitopes in the sense as defined by Jerne, but simply short, linear sequences within the unfolded protein which at most could correspond to cryptotopes as defined by Jerne. Peptides are of limited use in epitope mapping or the study of antigenic structures of native proteins. In fact, in most cases anti-peptide antibodies or unfoldons identified by them are devoid of any biological significance, and results from such epitope mapping studies can be extremely misleading.

Another frequently held misconception is that proteins display only a restricted small number of epitopes in a polyclonal response. Although some specificities may be immunodominant owing to immune regulation, the evidence that most, if not all, of the protein surface is antigenic has been considered previously in detail (Benjamin et al., 1984) and is an underlying assumption of our discussion.

In addition to the use of anti-peptide antibodies, other functional methods that measure some aspect of antibody–antigen interaction have been used in futile attempts to define protein epitopes. These include:

- absorption of sera with fragments of the antigen,
- antibody protection (of residues against chemical derivatization or peptide bonds against hydrolysis),
- competition tests among monoclonal antibodies,
- reactivity of anti-protein antibodies with variants of the antigen (which include naturally occurring antigenic variants and variants produced by site-directed mutagenesis or selection of escape mutants of viral antigens), and
- molecular modeling (from physical properties of antigens and from energetic analyses of complexes—an approach that has also been used to predict antigenic regions or residues).

Some of the above methods have been partially successful in identification of specific residues or general regions recognized by an antibody, but none has yet allowed the complete definition of an epitope. To date, only determination of the structure of crystals of antibody–antigen complexes has allowed complete definition.

Attempts to match the functional epitope and the structural epitope precisely have not yet been completely successful, although the disparity is not great. Functional assays may indicate that a smaller subset of residues is involved in binding energy, but this probably reflects a limitation on the number of antigenic variants available. In fact, for all the complexes, variants of only a small fraction of the contact residues have been tested. For example, competitive inhibition studies with evolutionary variants of lysozyme correctly identified critical residues in the HyHEL-5 epitope but did not predict its full extent. Similar methods were less successful in predicting the HyHEL-10 contact residues (Lavoie et al., 1990). In addition, all escape mutants that abolish binding of the NC41 antibody to neuraminidase do in fact correspond to residues within the structural epitope. To date, there are no published examples of a mutation of a residue within the structural epitope which does not affect binding.

Molecular modeling has suggested that a subset of residues within the structural epitope may contribute most of the binding energy, i.e., form an "energetic epitope," with surrounding residues allowing structural complementarity (Novotny, 1990). It should be noted that these energetically critical residues are generally not sequentially located. Nevertheless, mutation of the surrounding residues may significantly reduce or even abolish binding. Whether the alteration of binding by changing any given amino acid is due to loss of energy contributing to the interaction by that particular residue, or whether it represents introduction of a "pin" in the interface that interferes with the way the antibody "sits" on the antigen will require refinement of structures of complexes with mutant antigens.

In conclusion, the structural data establish that epitopes on native proteins consist of 15–22 residues in a discontinuous array. Energetic calculations suggest that a smaller subset of 5–6 of these residues contributes most of the binding energy, with the surrounding residues merely indulging in complementarity. It should be stressed that the residues proposed to contribute most of the binding energy are not arranged in a linear sequence but are scattered over the epitope surface; in no sense can they be considered equivalent to unfoldons identified with antisera against short peptides. Ultimately, definition of the precise relationship between the structure and the function of the epitope will require detailed kinetic and structural analysis of site-directed mutants of both antigen and antibody.

References

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986). Science *233*, 747–753.

Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlen, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., and Wilson, A. C. (1984). Annu. Rev. Immunol. *2*, 67–101.

Davies, D. R., Padlan, E. A., and Sheriff, S. (1990). Annu. Rev. Biochem. *59*, in press.

Goodman, J. W. (1989). In Antigenic Determinants and Immune Regulation, E. Sercarz, ed. (Basel: Karger), pp. 1–22.

Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. N., Sutcliffe, J. G., and Lerner, R. A. (1982). Cell *28*, 477–487.

Jemmerson, R., and Blankenfeld, R. (1989). Mol. Immunol. *26*, 301–307.

Jerne, N. K. (1960). Annu. Rev. Microbiol. *14*, 341–358.

Lavoie, T. B., Kam-Morgan, L. N. W., Mallett, C. P., Schilling, J. W., Prager, E. M., Wilson, A. C., and Smith-Gill, S. J. (1990). In Use of X-ray

Crystallography in the Design of Antiviral Agents, W. G. Laver and G. M. Air, eds. (New York: Academic Press), pp. 213–230.

Novotny, J. (1990). Mol. Immunol., in press.

Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., and Davies, D. R. (1989). Proc. Natl. Acad. Sci. USA. *86*, 5938–5942.

Sela, M. (1969). Science *166*, 1365.

Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., and Davies, D. R. (1987). Proc. Natl. Acad. Sci. USA *84*, 8075–8079.

Tulip, W. R., Varghese, J. N., Webster, R. G., Air, G. M., Laver, W. G., and Colman, P. M. (1990). Cold Spring Harbor Symp. Quant. Biol. *54*, 257–263.