Three-dimensional structure of an antibody–antigen complex

(immunoglobulins/epitope/x-ray crystallography/complementarity/lysozyme)

STEVEN SHERIFF*, ENID W. SILVERTON*, EDUARDO A. PADLAN*, GERSON H. COHEN*, SANDRA J. SMITH-GILL†, BARRY C. FINZEL‡§, AND DAVID R. DAVIES*

*Laboratory of Molecular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, Bethesda, MD 20892; †Laboratory of Genetics, National Cancer Institute, Bethesda, MD 20892; and §Genex Corporation, Gaithersburg, MD 20877

Contributed by David R. Davies, July 30, 1987

ABSTRACT We have determined the three-dimensional structure of two crystal forms of an antilysozyme Fab–lysozyme complex by x-ray crystallography. The epitope on lysozyme consists of three sequentially separated subsites, including one long, nearly continuous, site from Gly-41 through Tyr-53 and one from Gly-67 through Pro-70. Antibody residues interacting with lysozyme occur in each of the six complementarity-determining regions and also include one framework residue. Arg-45 and Arg-68 form a ridge on the surface of lysozyme, which binds in a groove on the antibody surface. Otherwise the surface of interaction between the two proteins is relatively flat, although it curls at the edges. The surface of interaction is approximately 26 × 19 Å. No water molecules are found in the interface. The positive charge on the two arginines is compensated by the negative charge of Glu-35 and Glu-50 from the heavy chain of the antibody. The backbone structure of the antigen, lysozyme, is mostly unperturbed, although there are some changes in the epitope region, most notably Pro-70. One side chain not in the epitope, Trp-63, undergoes a rotation of ~180° about the Cβ–Cγ bond. The Fab elbow bends in the two crystal forms differ by 7°.

Until recently knowledge of the structural aspects of antibody–antigen interactions has been based on the x-ray analysis of four Fab structures and on some complexes with haptens (1–5). Haptens were observed to bind in grooves or pockets in the combining sites of the New and McPC603 Fabs, and these occupied a small fraction of the total available area of these sites. When haptens bind to these Fabs, no large conformational change occurs. However, one cannot rule out the possibility that the behavior of antibodies would be different when they are bound to larger antigens, such as proteins. For example, the interaction with a much greater fraction of the combining site might in itself be sufficient to induce conformational changes in the antibody. As the interacting surfaces might not possess the grooves and pockets observed for haptens, but might resemble more closely the kind of surface observed in other protein–protein interfaces, where exclusion of bound water is believed to play a key role. For this reason we undertook several years ago to investigate the crystal structures of complexes of the Fabs of several monoclonal antibodies to hen egg white lysozyme complexed with the lysozyme (6). In this paper we report the analysis of two different crystal forms of one of these complexes.

The site on the lysozyme to which the antibody binds has been the subject of an extensive serological analysis (7) through a study of cross-reactivity with different avian lysozymes. The results of that analysis are in striking agreement with the crystal structure observations and will be discussed.

During the course of this analysis two reports of related x-ray studies of Fab–antigen complexes have appeared (8, 9), one being a description of another lysozyme–antilysozyme complex, although to a different epitope of the lysozyme, and the other describing a complex with the neuraminidase of influenza virus. The observations and conclusions from these two investigations differ in important ways from one another, and we describe below how our results can be related to them.

MATERIALS AND METHODS

Monoclonal antibody (mAb) HyHEL-5 and the Fab–lysozyme complex were prepared as described previously (6, 7). Crystals were grown by vapor diffusion against 20% (wt/vol) polyethylene glycol 3400 (Aldrich) in 0.1 M imidazole hydrochloride, pH 7.0, 10 mM spermine with an initial protein concentration of 7 mg/ml. The crystals grow polymorphically in space group P2₁2₁2₁, differing principally in the length of the b axis, which was observed to vary between crystals in an unpredictable manner between 65 Å and 75 Å.

One set of data on a crystal with cell dimensions of a = 54.9 Å, b = 65.2 Å, c = 78.6 Å, and β = 102.4° was collected at Genex (Gaithersburg, MD) using a single detector–single axis Nicolet–Xentronics (Madison, WI) area detector; 20,074 observations yielded 7565 unique reflections. Lorentz, polarization, and absorption corrections were applied (10), and the different frames were scaled together giving an overall merging R_sym of 0.044, where R_sym = ΣhkI₁ − I̅I/ΣhkI₁. Greater than 90% of the theoretical data were observed to 4.5 Å spacings, greater than 80% from 4.5 Å to 4.0 Å spacings, and about 40% from 4.0 Å to 3.0 Å spacings.

A second data set was collected on a crystal with cell dimensions of a = 54.8 Å, b = 74.8 Å, c = 79.0 Å, and β = 101.8° at the University of California, San Diego, using the Mark II multiwire detector system with two detectors (11); 38,689 reflections were collected from one crystal, of which 15,673 were unique. Lorentz, polarization, and absorption corrections were applied, and the different frames were scaled together giving an overall merging R_sym of 0.044. Of the 15,673 unique reflections, 15,166 are within 2.66 Å resolution (86.3% of the theoretically observable); there are an additional 507 reflections between 2.66 and 2.54 Å resolution (18.7% of the theoretically observable).

The structure was determined by molecular replacement using the program package assembled by Fitzgerald (12). Three probes were used: (i) tetragonal lysozyme (2LYZ) deposited by R. Diamond in the Protein Data Bank (13); (ii) C_L + C_H1 of the McPC603 Fab (4); and (iii) V_L + V_H of the

Abbreviations: CDR, complementarity-determining region; CDR1, CDR2, and CDR3 for the light chain are referred to as L1, L2, and L3, and for the heavy chain as H1, H2, and H3; mAb, monoclonal antibody; C, constant region; V, variable region.

‡Present address: Central Research and Development, E. I. DuPont Experimental Station ES 228/316B, Wilmington, DE 19898.
McPc603 Fab with the following residues from the model: \( V_l \) -27C-31 and 91-95; and \( V_h \) -30-31, 52B-54, 61-64, 96-100, and 101. Residue numbering in Fabs throughout this paper follows Kabat et al. (14); C and V represent constant and variable regions, respectively, and L and H refer to light and heavy chain, respectively.

We oriented McPc603 Fab so that the axis of the elbow was parallel to the \( z \) axis, which allowed us to observe most of the difference in elbow bend directly in the \( y \) rotation function angle (15). The fast-rotation function (16) was used with 10- to 4-A resolution data and a radius of integration of 24 \( \AA \). The rotation function of Lattman and Love (17) was used to refine the position of the peak for each probe. The Crowther-Blow (18) translation function was used with an interatomic distance of 3.0 \( \AA \) and a step size of 0.02 unit cell lengths in \( a \) and \( c \) to determine the \( x \), \( z \) translations. We also used the program BRUTE, written by M. Fujinaga and R. Read, University of Alberta, with 5-to 4-A resolution data and 1-A step size to determine \( x \), \( z \) translations.

Where the two methods agreed, we used BRUTE to hold one or two probes stationary and search for the translation of the second or third probe to solve the problem of relative origins in space group \( P2_1 \). Where the methods disagreed, we used the program PRECOR/CORELS (20) to hold one or two probes stationary and search for the translation of the second or third probe to solve the problem of relative origins in space group \( P2_1 \).

We then divided the complex into five domains: lysozyme, \( V_l \), \( V_h \), \( C_l \), and \( C_h \). At this point we replaced the \( V_l \) domain of McPc603 with that of antibody J539 (5), because it shares 75% sequence identity with the murine IgGl and \( V \) domains of McPc603, which is a murine IgGl, and shares 60% sequence identity with the murine IgGl \( C \) domain of mAb HyHEL-5. At this stage of refinement for both crystal forms the \( R \) was 0.42 for 10- to 6-A resolution data.

At this point we concentrated on the long \( b \)-axis form because the data set extended to higher resolution. We removed from the model the side chains of amino acids that were not identical to the mAb HyHEL-5 sequence (ref. 21 and A. B. Hartmann, C. P. Mallett, and S. J. S. G., unpublished work) and also omitted entire residues of the following complementarity-determining regions (CDRs): L3 (residues 91-95), H2 (residues 52B-54), and H3 (residues 97-100); (see abbreviations footnote). We refined the model against data from 10- to 2.5-A resolution using the stereochemically restrained least-squares refinement package PROTON/PROLSQ (22). We first refined only the positional parameters using an overall isotropic B until the \( R \) was 0.344. We then used FRODO to rebuild the model. In the electron-density map at this stage there was excellent density for the deleted L3 and moderately good density for the deleted H2 and H3. Also, in most cases, electron density was apparent for side chains that had been omitted from the model. Following this, we refined the model adding individual isotropic B factors until the \( R \) was 0.270, and then we examined the model with molecular graphics and rebuilt parts of the model, especially H3. We once again refined until the \( R \) was 0.249. At this stage we included an overall anisotropic B (23), which had values of \( \Delta B_{11} = -5.89 \), \( \Delta B_{12} = 2.15 \), \( \Delta B_{22} = 8.44 \), and \( \Delta B_{33} = -2.55 \) and which correlates with the variable \( b \)-axis length, suggesting the lattice contacts along the \( b \) axis are weak. We then omitted all of the interacting residues in lysozyme, refined the model, calculated an electron-density map, and examined it to determine whether our interpretation of the position of these residues was correct. We did the same thing with the interacting residues on the Fab. After making minor changes to the structure we refined again, yielding \( R = 0.245 \) with an rms deviation from ideal bond lengths of 0.012 \( \AA \). We have deposited the coordinates from this stage of the refinement in the Protein Data Bank (13).

We have estimated the rms positional error to be 0.40 \( \AA \) by the method of Luzzati (24). In describing the results, hydrogen bonds and salt links were limited to pairs of appropriate atomic groups at an interatomic distance of \( < 3.0 \AA \). Van der Waals contacts distance were defined as in Sheriff et al. (25). Contacting surface area was calculated with program MS (26) using a probe radius of 1.7 \( \AA \) and standard van der Waals radii (27).

RESULTS

Fig. 1a shows the \( C \) skeleton of the HyHEL-5 Fab-lysozyme complex. \( V_h \) and \( V_l \), and \( C_h \) and \( C_l \) adopt the canonical relationships observed in other Fabs. The main difference between the two crystal forms is the elbow bend of the HyHEL-5 Fab, which is 161° in the long \( b \)-axis form and 154° in the short \( b \)-axis form.

The contact between the antibody-combining site and the lysozyme epitope is extensive and involves many residues. The calculated buried surface (solvent inaccessible) area is about 750 \( \AA^2 \) on the surface of both the Fab and lysozyme (14% of the surface). The interaction between the two proteins is very tight, and there are no water molecules between the combining sites of the Fab and lysozyme. The current model contains three salt links and ten hydrogen bonds.

Gl-50 (H2) forms salt links to both Arg-45 and Arg-68 of lysozyme, and Glu-55 (H1) forms a salt link to Arg-68. There are 74 van der Waals contacts.

The epitope on lysozyme consists of three oligopeptide segments (Fig. 1b). The first consists of Gin-41, Thr-43, Asn-44, Arg-45, Asn-46, Thr-47, Asp-48, Gly-49, and Tyr-53, which are in contact, and Thr-51 and Asp-52, which are partly buried by the interaction. This segment is essentially one long continuous subsite, which consists of two \( \beta \)-strands connected by a bend involving residues 46 through 49 (28). The second segment consists of the contacting residues Gly-67, Arg-68, Thr-69, and Pro-70; and the partly buried residues Asn-65, Asp-66, Gly-71, and Ser-72. This second segment has the form of a rambling loop and contains part of an extensively studied, disulfide-linked antigenic peptide that elicits antibodies that could cross-react with native lysozyme (29). Arg-68 in this subsite has been identified as a "critical" residue to the epitope (7). The third segment consists of the directly interacting Leu site and the partly buried residues Pro-79, Ser-81, and Ser-85. Arg-61, which is in none of the segments, is also partly buried upon complex formation. The surface of the epitope is extensive (23 \( \AA \) between the most distant \( C \) atoms) and relatively flat except for a protruding ridge made up of the side chains of Arg-45 and Arg-68, and curling back at the edges (Fig. 1d).

The antibody-combining site involves residues from all six CDRs (30). Each of these CDRs contributes at least one residue to the interaction with lysozyme, and most contribute several residues (Fig. 1c). Trp-47, which is considered part of the heavy chain framework, also interacts with lysozyme. The other interacting residues are Asn-31, Tyr-32, Asp-40, Trp-91, Gly-92, Arg-93, and Pro-95 from the light chain; and Trp-47, Glu-35, Gly-36, and Arg-68 from the heavy chain. Additional residues that are at least partly buried by the interaction but do not directly contact are Ser-29, Val-30, Tyr-34, and Arg-46 (in...
framework region 2) from the light chain and Ser-30, Asp-31, Tyr-32, Leu-52, Gly-55, Asn-96, and Asp-98 from the heavy chain. Only 30% of the residues in the CDRs are actually in contact with lysozyme. If one adds the residues that are at least partly buried, this fraction rises to 48%. The surface of interaction is quite broad and extensive with the C" atoms of interacting combining site residues separated by as much as 28 A. There is a groove on the surface of the antibody running from L3 to H3 and between Trp-91 of L3 and Trp-33 of H1 (Fig. 1d). Arg-45 and Arg-68 of lysozyme fit into this groove, placing them in position to form salt links to Glu-50 (H2) and Glu-35 (H1).

We examined the lysozyme and Fab structures for indication at this degree of refinement of any significant conformational change that may have occurred as a result of their association. The structure of hen egg lysozyme has been determined in four crystal forms, thus providing a database for comparing the structure of lysozyme in different environments. The tetragonal lysozyme coordinates used in this comparison (D. C. Phillips, personal communication) differed slightly from those used for structure determination (rms difference = 0.32 A for 516 main-chain atoms). We superimposed the tetragonal lysozyme coordinates onto the lysozyme in our complex, using all main-chain atoms, both including and excluding atoms from residues that are involved in interactions with the antibody-combining site on the Fab. Qualitatively we see the same behavior with both procedures, and we report numbers calculated when residues were excluded (Fig. 1e). For these lysozymes, the backbone structures are nearly identical (rms difference = 0.48 A for

![Figure 1](image_url)
516 atoms). However, we do see some changes in the region of the epitope (rms difference = 0.64 Å for 96 atoms). In particular, the C\textsuperscript{\beta} of Pro-70 has moved about 1.7 Å, which can be accounted for by a hydrogen bond of the carbonyl oxygen and the hydroxyl of Tyr-97 (H3).

There are also some changes in the side-chain atoms (rms difference = 1.22 Å for 429 atoms). The side chains of residues not involved in the epitope (excluding Trp-63) show an rms difference of 1.12 Å for 420 atoms. Comparing side chains of residues involved in the epitope gives an rms difference of 1.20 Å for 55 atoms. The largest side-chain difference is Trp-63, which is not in the epitope and which appears to have flipped 180° around the C\textsuperscript{\beta}—C\textsuperscript{\gamma} bond. Although the current electron-density maps are not strongly favoring movement, there is some ambiguity because there is a tail of electron density that points, more or less, in the direction of the original side-chain conformation. Preliminary refinement of the short b-axis form strongly suggests the movement of the Trp-63 side chain. Many of the other large differences are in residues that are part of the epitope. In particular, the side chains of Arg-45 and Arg-68 in the tetragonal crystal form, the triclinic crystal form, and the complex with the antibody are within hydrogen-bonding distance of one another, but in none of the three are the side chains in the same position relative to the backbone.

To ascertain whether any changes have taken place in the Fab in the complex, we need crystals of the uncomplexed Fab, but so far we have been unable to grow such crystals large enough for diffraction studies. However, the fact that the molecular replacement technique worked so well shows that the positions of Cl\textsubscript{\beta} relative to C\textsubscript{\gamma1} and of V\textsubscript{H} relative to V\textsubscript{L} are basically unchanged from the “canonical” structures found in Fab fragments from myeloma proteins. This result differs from the result of Colman \textit{et al.} (9), who report that V\textsubscript{L} and V\textsubscript{H} form a non-canonical pairing in their antineuraminidase–neuraminidase complex.

**DISCUSSION**

Our results confirm the identification of the epitope on lysozyme for HyHEL-5 based on serological studies (7). In those studies, epitope mapping was accomplished by analyzing the cross-reactivity of homologous proteins in the antibody–antigen reaction. The limited sequence variation among these cross-reacting proteins was utilized to pinpoint the residues that were important in the particular binding interaction. The present crystallographic results provide strong evidence for the validity of epitope mapping by serological techniques.

The HyHEL-5–lysozyme complex described here and the two other antibody–antigen complexes whose structures have been reported share some common features. For example, in all three cases, the interaction involves all of the CDRs of the antibody. The surface area of the combining site of the antibody that is buried upon complexing with antibody is 750 Å\textsuperscript{2} for HyHEL-5, 690 Å\textsuperscript{2} for the antilysozyme Dl.3 Fab (8), and not specified for the antineuraminidase NC41 Fab (9), where the third CDRs of both light and heavy chains have not yet been completely modeled.

Although the areas of interaction are comparable for the HyHEL-5 and Dl.3 antibody, the association constant is greater for HyHEL-5 than for Dl.3 antibody (2.5 × 10\textsuperscript{8} (M. E. Denton and H. A. Scheraga, personal communication) vs. 4.5 × 10\textsuperscript{8} (8)). This indicates that the area of interaction alone cannot be the total determinant of the attraction between the antibody and antigen; undoubtedly, other factors such as electrostatic and hydrogen bonding play a role. In this connection, the favorable electrostatic interactions involving Arg-45 and Arg-68 of lysozyme and Glu-35 (H1) and Glu-50 (H2) of the Fab, and the shape complementarity between the Arg-45, Arg-68 side-chain ridge on lysozyme and the groove on the Fab are relevant. In contrast, there are no electrostatic interactions between Dl.3 antibody and lysozyme (8). We note that when Arg-68 becomes a lysine, as it does in bobwhite quail, the HyHEL-5 antibody binds with less avidity by a factor of 10\textsuperscript{4} (7), so that shape and hydrogen-bonding capacity of the charged group are also crucial. In antibody Dl.3, H3 plays a dominant role in the antibody–antigen interaction (8). However, in HyHEL-5, H2 and L3 play the dominant role contributing six and four residues, respectively, and approximately equal surface areas to the interaction. CDR 1 of the light chain (L1), H1, and H3 play a lesser role by each contributing two residues and over 50–60% of the surface area of H2 and L3.

The lysozyme epitope for HyHEL-5 consists principally of two continuous segments of polypeptide chain, residues 41–53 and 65–72. The lysozyme epitope for antibody Dl.3 again involves two continuous segments of the protein, residues 18–27 and 116–129 (8). This similarity is remarkable, but it is probably coincidental because the epitope for NC41 Fab on the neuraminidase involves four segments, 368–370, 400–403, 430–434, and parts of 325–350 (9). It is interesting that the two continuous segments in the lysozyme epitope for mAb HyHEL-5 have been reported to have high mobility in several lysozyme structures (31). Conformational changes in the Fab that result from antigen binding cannot be quantitated because in none of the three cases so far reported has the structure of the uncomplexed Fab been analyzed. It should be noted that antibodies frequently have large insertions into the CDRs, and these sometimes project into the solvent and cannot be localized with certainty by x-ray diffraction. L1 in McPC603 Fab is an example of this (4). In any interaction involving McPC603 Fab with a large antigen, it is likely that such a loop would move. However, Colman \textit{et al.} (9) have reported a difference in the relationship of V\textsubscript{L} to V\textsubscript{H} in NC41 Fab. We have duplicated the calculations in Table 1 of Colman \textit{et al.} (9) with results that are essentially in agreement with theirs and extended the table to include Fabs J539 and HyHEL-5. We find that the relative disposition of the two variable domains in the NC41 Fab does lie at the extreme of the values observed. Whether this is the result of binding to antigen or the result of the interaction of hypervariable residues in the interface cannot be determined at present. We should point out, however, that the NC41 Fab–neuraminidase complex has been elucidated to only 3.0-Å resolution and has been subjected to only preliminary refinement (R = 0.35). These conclusions need to be confirmed by further refinement.

Amit \textit{et al.} (8) have observed that no large conformational changes have occurred in lysozyme upon binding to the Dl.3 Fab. Over most of the molecule we find essentially the same results, although some changes do occur in both the backbone and the side chains. In particular, the flipping of the Trp-63 side chain, if confirmed by further refinement, together with movement of Pro-70 are the most notable of these changes. Colman \textit{et al.} (9) have also reported a few significant differences between the structure of the neuraminidase alone and that in the complex with NC41 Fab, although here, too, we must await confirmation from further refinement. It is apparent that some deformation of the antigen can occur, especially when the epitope is in a flexible part of the structure, as in HyHEL-5–lysozyme complex. Indeed, flexibility has been implicated in the antigenicity of proteins (32, 33) and could aid in the binding of antibody to antigen by allowing the latter to complement more closely the structure of the antibody–antigen complex.

It has been hypothesized that changes in the elbow bend may signal antigen binding (34). The different crystal forms of the HyHEL-5–lysozyme complex are the first example of the same antibody–antigen complex showing different elbow-
bend angles. The values observed here lie in the middle of the spectrum of observed elbow bends ranging from 133° to
≈180° (35). There appears to be no correlation between observed values of elbow bends for different Fab's and
binding to hapten or antigen. It is therefore likely that this range of values is simply an indication of flexibility of this part
of the antibody molecule.

In conclusion, the results observed for the binding of antibodies to protein antigens have many features in common
with the binding of haptens. For example, it is clear that
the much greater area of the complementary surfaces that are
brought into contact with consequent exclusion of water
13. Bernstein, F. C .. Koetzle . T. F ..