Structural Insights into the Evolution of an Antibody Combining Site

Gary J. Wedemayer, Phillip A. Patten,* Leo H. Wang, Peter G. Schultz,† Raymond C. Stevens

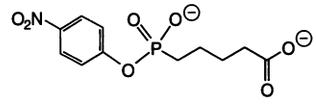
The crystal structures of a germline antibody Fab fragment and its complex with hapten have been solved at 2.1 Å resolution. These structures are compared with the corresponding crystal structures of the affinity-matured antibody, 48G7, which has a 30,000 times higher affinity for hapten as a result of nine replacement somatic mutations. Significant changes in the configuration of the combining site occur upon binding of hapten to the germline antibody, whereas hapten binds to the mature antibody by a lock-and-key fit mechanism. The reorganization of the combining site that was nucleated by hapten binding is further optimized by somatic mutations that occur up to 15 Å from bound hapten. These results suggest that the binding potential of the primary antibody repertoire may be significantly expanded by the ability of germline antibodies to adopt more than one combining-site configuration, with both antigen binding and somatic mutation stabilizing the configuration with optimal hapten complementarity.

One mechanism whereby the immune system recognizes and distinguishes foreign antigens is by generating a large and diverse repertoire of antibody molecules. This diversity protects organisms from a wide range of pathogens and toxic agents and has been exploited to produce high affinity–selective receptors for use as diagnostics, molecular probes, and therapeutic agents. With the proper chemical instruction, the molecular diversity of the immune system can be brought into the service of chemistry as a source of selective antibody catalysts (*1*); it has also served as the inspiration for synthetic combinatorial libraries of biomolecules, synthetic organic molecules and solid-state compounds (*2*).

Many theories have been put forth to explain the ability of antibodies to recognize a seemingly unlimited number of antigens. Early theories held that the antigen served as a template for the biosynthesis of a complementary combining site (*3*). Alternatively, antigen could serve as a template for the folding of one of many possible configurations of a single polypeptide chain (*4*). These instructional theories have since been replaced by the clonal selection theory in which each lymphocyte produces a distinct antibody and clones are selected on the basis of the affinity of antibody for antigen (*5*). Diversity in the germline antibody population is generated by the combinatorial association of V, D, and J gene segments with additional junctional diversity occurring at the V_L–J_L, V_H–D, and D–J_H joining regions because of imprecise joining and addition of N region nucleotides (*6*). Somatic mutation, which alters bases throughout the sequences encoding the variable region, provides further diversity and leads to increased affinity and specificity as the immune response proceeds (*6, 7*). Although genetic and biochemical studies have revealed the nature and origin of the sequence diversity of antibodies, the structural basis for the transformation of this sequence diversity into tailor-made high affinity receptors is less well understood (*8, 9*).

As part of our efforts to explore the immunological evolution of antibody catalysis, we have been characterizing the biophysical, kinetic, and structural properties of a series of catalytic antibodies, their germline precursors, and related mutants (*10–13*). In this article, we compare the high resolution x-ray crystal structures of the Fab fragments of the esterolytic antibody 48G7 (*10, 14*) and its germline precursor. In each case, structures were solved for both the Fab fragment and the complex of Fab with the nitrophenyl phosphonate hapten **1**. An analysis of these structures reveals that significant structural changes occur in the variable region in response to hapten binding to the germline antibody and on affinity maturation of the germline antibody to mature immunoglobulin. These studies provide important insights into the molecular basis of the immune response, which may also bear on other combinatorial systems for evolving new function.

Structure determination. The germline precursor to antibody 48G7 binds nitrophenyl phosphonate transition-state analog: **1** (scheme 1) with a dissociation constant



1

(K_d) of 135 μM (*10, 14*). Antibody 48G7, which differs by six amino acid changes in the heavy chain (Glu_G^{H42} → Lys, Gly_G^{H55} → Val, Asn_G^{H56} → Asp, Gly_G^{H65} → Asp, Asn_G^{H76} → Lys, and Ala^{H78} → Thr) and three changes in the light chain (Ser_G^{L30} → Asn, Ser_G^{L34} → Gly, and Asp_G^{L55} → His), binds hapten **1** with a K_d of 4.5 nM (*10*). This 30,000 times higher affinity results primarily from a decrease in the rate of antibody–hapten **1** dissociation and has been correlated with an appreciable increase in the catalytic efficiency of the an-

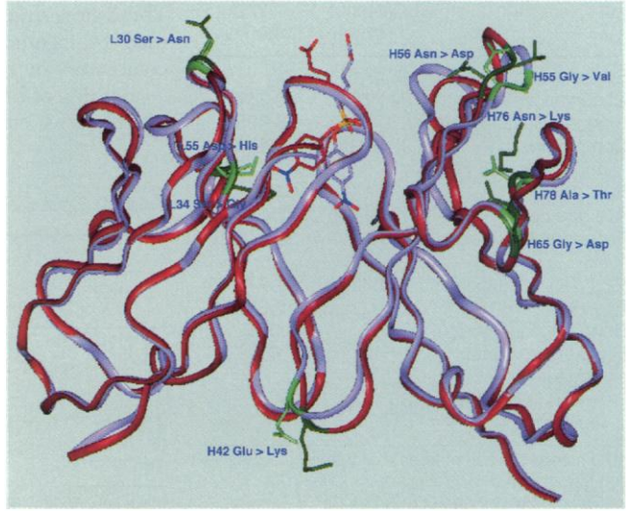


Fig. 1. Ribbon superpositions of the variable regions of the germline Fab-hapten complex (light purple) and mature Fab-hapten complex (dark red). The aliphatic linker used to conjugate the hapten to the carrier protein can be seen extending toward the top of the figure. The side chains of the somatic mutation sites are indicated in light green (germ line) and dark green (mature) (Ser^{L30} → Asn, Ser^{L34} → Gly, Asp^{L55} → His, Glu^{H42} → Lys, Gly^{H55} → Val, Asn^{H56} → Asp, Gly^{H65} → Asp, Asn^{H76} → Lys, Ala^{H78} → Thr).

The authors are in the Department of Chemistry, University of California, Berkeley, CA 94720, USA and at the Lawrence Livermore National Laboratory, Berkeley; P. G. Schultz is also with the Howard Hughes Medical Institute.

*Present address: Maxygen, Inc., 3410 Central Expressway, Santa Clara, CA 94501, USA.
†To whom correspondence should be addressed.

tibody (*10*). The x-ray crystal structures of the Fab fragment of 48G7 and its complex with transition-state analog **1** (Fab-hapten complex at 2.0 Å resolution, Fab at 2.7 Å resolution) indicate that none of the nine residues in which somatic mutations had been fixed directly contact the hapten in the structure of the mature antibody (Fig. 1) (*10*).

In order to better understand the structural basis for the large change in affinity associated with these somatic mutations, the x-ray crystal structures of the germline Fab fragment and its complex with hapten **1** were both determined at 2.1 Å resolution by molecular replacement with the affinity-matured Fab structure as a starting model (*15, 16*). The final statistics for the structures are listed in Table 1. Nearly identical crystallization conditions were used for the germline and mature antibodies. Both the Fab fragment of 48G7 and its complex with hapten **1** crystallized in the space group C2, whereas the germline Fab-hapten complex crystallized in the space group $P2_1$ and the unliganded germline Fab crystallized in the orthorhombic space group $P2_12_12_1$.

Since the germline Fab binds hapten **1** with much lower affinity than the affinity-matured Fab, a 50-fold molar excess of hapten **1** was used to crystallize the germline Fab-hapten complex. The electron density of the phosphonate group is well defined in both the germline and mature complexes. However, the electron densities of the nitrophenyl and aliphatic linker groups of the hapten are of higher quality in the mature complex than in the germline complex. For example, although the electron density of the nitroaryl ring is present in the germline structure, it is not as highly defined as in the mature complex where a hole in the aryl ring can be visualized. A similar trend was observed for the aliphatic linker of hapten **1**. No well-defined water or buffer molecules were observed in the oxyanion hole for the unliganded forms of either the germline or affinity-matured Fabs. The closest water molecule present in the antigen binding site of the germline Fab is 3.7 Å away from His^{H35} and oriented away from Tyr^{H33} and Arg^{L96}.

Structural consequences of hapten binding. Comparison of the structure of the unliganded Fab with the Fab–hapten **1** complex for the affinity-matured antibody, 48G7, reveals that very few structural changes occur upon binding of hapten (Table 2) (*10*). The root-mean-square (rms) deviation for all the Cα atoms of the variable region is 0.39 Å. The relative domain association (*17*) between the framework regions of V_L and V_H changes by only 0.44° on hapten binding. There are also no significant changes either in the packing of V_H and V_L or in the positions of the key active site residues. These include Arg^{L96}, Tyr^{H33}, and His^{H35}, which hydrogen bond to the phosphonyl oxygens of hapten **1**; Trp^{H47}, Ser^{L93}, Tyr^{L94}, and Arg^{H50}, which fix the orientation of these three oxyanion binding residues; and Tyr^{H98}, Tyr^{H99}, Tyr^{L91}, and Leu^{L89}, which form van der Waals contacts with the hapten.

In contrast to the lock-and-key fit (*18*) of hapten **1** to 48G7, binding of hapten to the germline antibody leads to significant structural changes. With respect to the liganded and unliganded germline antibody Fab structures, the rms deviation for all the Cα atoms of the variable region is 0.61 Å and the relative domain association in the germline structure changes by 4.6° (Table 2) (*17*). Whereas there are 16 hydrogen bonding, electrostatic, and van der Waals interactions between the V_H and V_L domains in the free germline Fab, there are 26 such interactions in the germline Fab-hapten **1** complex. These gross structural changes are accompanied by significant reorganization in the combining site residues (Fig. 2).

Binding of hapten leads to repositioning of the active site residue, Tyr^{H33}, and the formation of a network of three hydrogen bonds between the side chains of Tyr^{H33}, Arg^{H50}, and Tyr^{L94} (Fig. 2B). These interactions, which are reinforced by the $Asn^{H56} \rightarrow Asp$ and $Gly^{H55} \rightarrow Val$ somatic mutations, play a key role in the formation of the oxyanion hole in the mature antibody by fixing the position of the Tyr^{H33} hydroxyl group (in contrast, the oxyanion hole in most proteases involves a fixed backbone hydrogen bond). Furthermore, the aromatic side chains of the active site residues Tyr^{H33} and Tyr^{H98} move 5 Å closer together into a T-stack arrangement, resulting in packing interactions between Tyr^{98H} and the aliphatic linker of the hapten (Fig. 2A). This change in configuration is accompanied by the formation of a π-stacking interaction between Tyr^{H98} and Tyr^{H99} and a π-cation interaction between the aryl ring of Tyr^{H99} and the Arg^{L46} side chain (*19*): Arg^{L46} is oriented by Asp^{L55}, which is a site of somatic mutation (Fig. 2A).

Although conformational changes have been observed on binding of antigen to a number of affinity-matured antibodies (*20*), the important point here is that for the 48G7 system, the structural changes that occur in the germline-hapten complex become preorganized in the combining site of the mature antibody. It is therefore of interest to determine whether a greater degree of conformational flexibility exists in the germline structures of these other antibody-antigen complexes (*20*).

Structural effects of somatic mutations. The crystal structure of the 48G7–hapten **1** complex indicates that the reorganization of the combining site, which occurred on binding of hapten **1** in the germline antibody, is further optimized by affinity maturation (Fig. 3). The phosphonate moiety of the hapten, which appears to be a major binding determinant, is located in similar positions in both liganded structures and forms hydrogen bonds to Arg^{L96} and His^{H35} (Figs. 2C and 3C). In addition, a new hydrogen bond is formed in the mature antibody between the phosphonyl group of the hapten and the hydroxyl group of Tyr^{H33}. The side chain orientation of Tyr^{H33} is fixed by further elab-

Table 1. Data collection and refinement statistics.

	Germline Fab	
	Without hapten	With hapten
Space group	$P2_12_12_1$	$P2_1$
Observations (*N*)	94,282	80,267
Unique reflections (*N*)	27,752	23,005
Completeness to refined resolution	97.3 (91.2)	99.5 (98.1)
R_{sym} (*I*) (%)	8.0 (14.4)	6.9 (20.4)
Average $I/\sigma I$	15.7 (7.2)	15.0 (5.2)
Refinement resolution (Å)	20.0–2.1	20.0–2.1
R_{cryst} (%)	21.0 (21.4)	21.7 (24.6)
R_{free} (%)	25.9 (25.8)	25.8 (28.9)
Bond lengths rmsd (Å)	0.008	0.005
Bond angles rmsd (degees)	1.542	1.613
Residues in disallowed regions†	3	1

*The numbers in parentheses are for the highest resolution bin used in data processing or refinement (2.17 to 2.1 Å). †According to PROCHECK (*15*).

Table 2. Root-mean-square differences (Å) between germline (g) and mature (m) antibody Fabs.

Comparison*	V_L		V_H		Both chains		V_LV_H rotation
	Cα	All	Cα	All	Cα	All	
g− to g+	0.38	0.66	0.51	0.99	0.60	0.94	4.60°
g− to m−	0.46	0.92	0.60	1.05	0.78	1.17	6.89°
g− to m+	0.41	0.78	0.30	0.65	0.77	1.11	6.94°
g+ to m−	0.41	0.84	0.66	1.15	0.60	1.05	3.63°
g+ to m+	0.40	0.79	0.65	1.02	0.59	0.95	3.51°
m− to m+	0.30	0.65	0.45	0.84	0.38	0.75	0.44°

*g−, germline without hapten; g+, germline with hapten; m−, 48G7 without hapten; m+, 48G7 with hapten.

oration of the hydrogen bond network between Tyr^{H33}, Tyr^{L94}, and Arg^{H50} that was formed upon hapten binding to the germline antibody (Fig. 3B). The guanidinium group of Arg^{H50}, which bridges Tyr^{H33} and Tyr^{L94}, is oriented by a salt bridge to the side chain carboxylate group of the somatically mutated residue, Asp^{H56}. This interaction is stabilized by somatic mutation of the adjacent Gly^{H55} to Val, changing this loop to a noncanonical conformation (*21*). As a result the backbone is altered in this loop from a class 4 six-residue β-hairpin turn to a class 4 four-residue turn (Fig. 3B) (*22*). This change in backbone conformation also leads to two new salt bridges between the carboxylate group of Asp^{H56} and the ε-amino group of Lys^{H58}, which serve to further reinforce the hydrogen bond network.

Somatic mutation of $Ser^{L34} \rightarrow$ Gly leads to binding of the nitroaryl ring of hapten in a well-defined geometry in 48G7 (unlike the situation in the germline antibody), in which it interacts with the side chains of Tyr^{H99}, Tyr^{L91}, and Leu^{L89} (Fig. 3D) (mutation of $Gly^{L34} \rightarrow$ Ser in 48G7 results in a five times lesser affinity for hapten). This presumably is a result of the removal of an otherwise repulsive steric interaction between the nitro group of hapten **1** and the side chain of Ser^{L34}, which would protrude into the binding site region of the mature Fab. The interactions between Tyr^{H98}, Tyr^{H99}, and Tyr^{L91} and hapten that are induced by binding of hapten to the germline antibody are also further optimized by somatic mutation (Fig. 3A). The packing interactions between the side chain of Tyr^{H98} and both the aliphatic linker of the hapten and side chain of Tyr^{H33} increase. The aryl ring of Tyr^{H98} now forms a T-stack interaction with that of Tyr^{H99} (versus a π-stack in the germline–hapten **1** complex) (*17*). This reorganization also leads to improved packing interactions between Tyr^{H99}, Tyr^{L91}, and the nitroaryl ring of the hapten, but results in the loss of the side chain interaction between Arg^{L46} and Tyr^{H99} (Fig. 3A). This latter change is facilitated by the somatic mutation of $Asp^{L55} \rightarrow$ His, which removes the salt bridge between Asp^{L55} and Arg^{L46}; instead, the imidazole ring of His^{L55} hydrogen bonds to the main chain carbonyl of Ser^{L56}.

The configuration of CDR H1 (complementarity-determining region 1 of the heavy chain), which contains the active site residues His^{H35} and Tyr^{H33}, is also affected by the somatic mutations $Ala^{H78} \rightarrow$ Thr and $Asn^{H76} \rightarrow$ Lys, which pack against this CDR (Fig. 3E). In the germline antibody, the carboxamide side chain of Asn^{L76} hydrogen-bonds to the main-chain carbonyl groups of the framework residues Phe^{H27} and Ala^{H24}. The mutation of $Asn^{H76} \rightarrow$ Lys removes this interaction and allows the H25 to H32 region to be more flexible. The side chain of Lys^{H76} in the mature antibody hydrogen bonds to the backbone carbonyl group of Thr^{H73}. Somatic mutation of $Ala^{H78} \rightarrow$ Thr in 48G7 results in packing interactions with the side chains of Met^{H34}, Trp^{H36}, and Ile^{H51}, which assist in stabilizing the configuration of the critical active site residues H33 to H35 in CDR H1 (Fig. 3E).

The roles of the other somatic mutations in optimizing hapten affinity are less clear (Fig. 1). The somatic mutation of $Ser^{L30} \rightarrow$ Asn allows an additional hydrogen bond to be made between the carboxamide side chain and main-chain carbonyl group, perhaps stabilizing the turn at this site. The mutation of $Gly^{H65} \rightarrow$ Asp in the turn at the base of CDR H2 is correlated with phi and psi values for this residue that are in the disallowed region of the Ramachandran plot for the affinity-matured antibody (*23*). It is likely that this change plays a role in the observed alteration of CDR H2 from a canonical (germline) to a noncanonical (48G7) conformation. In the germline structure, Gly^{H65} may be a pivotal point from which CDR H2 moves in response to hapten binding. The somatic mutation of $Glu^{H42} \rightarrow$ Lys is located at the bottom of the turn in the framework region 2. In the structure of the germline–hapten complex, the whole loop containing residue H42 is puckered relative to that in other antibody structures. This seems to be a consequence of crystal packing in the germline-hapten complex because H40-H45 interacts with a symmetry-related antibody molecule. This conclusion is further supported by the fact that the loop in the germline structure without hapten has a conformation similar to that in the two affinity-matured antibody structures, which do not have this symmetry-related contact. The somatic mutations Asp^{H65} and Lys^{H42} may be neutral with regards to hapten binding, or they may affect the overall sta-

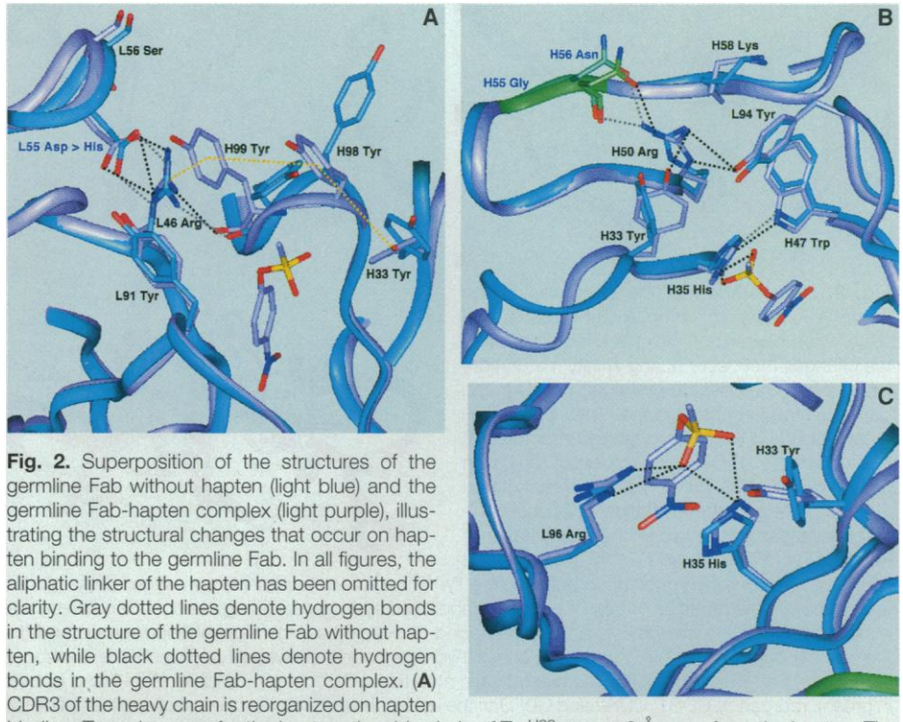


Fig. 2. Superposition of the structures of the germline Fab without hapten (light blue) and the germline Fab-hapten complex (light purple), illustrating the structural changes that occur on hapten binding to the germline Fab. In all figures, the aliphatic linker of the hapten has been omitted for clarity. Gray dotted lines denote hydrogen bonds in the structure of the germline Fab without hapten, while black dotted lines denote hydrogen bonds in the germline Fab-hapten complex. (**A**) CDR3 of the heavy chain is reorganized on hapten binding. To make room for the hapten, the side chain of Tyr^{H99} moves 6 Å away from the hapten. The side chain of Tyr^{H98} moves 8.3 Å and inserts between Tyr^{H99} and Tyr^{H33}, and Tyr^{H33} moves toward the phosphonate group. These movements establish a π-cation interaction between the side chains of Arg^{L46} and Tyr^{H99}, a π-π interaction between the aryl groups of Tyr^{H99} and Tyr^{H98}, and a T-stack interaction between the aryl rings of Tyr^{H98} and Tyr^{H33} (yellow dotted lines). In addition, the Arg^{L46} side chain is stabilized by salt bridges to the Asp^{L55} carboxylate group and to the Tyr^{H99} main chain carbonyl group. (**B**) The interactions between residues in CDR1, CDR2, and CDR3 of the heavy chain in the germline Fab structures. The side chain of Arg^{H50} forms hydrogen bonds to the hydroxyl groups of Tyr^{H33} and Tyr^{L94} upon hapten binding. The guanidinium group of Arg^{H50} is positioned by a hydrogen bond with Asn^{H56}. Although Tyr^{H33} forms one hydrogen bond to Arg^{H50}, it does not interact directly with either Tyr^{L94} or the bound hapten, nor does Lys^{H58} interact with residue H56 (*cf.* Fig. 3B). (**C**) Closeup of the combining site showing the orientations of the residues directly involved in hapten binding in the germline-hapten complex His^{H35}, Tyr^{H33}, and Arg^{L96}. All four hydrogen bonds are directed to the oxygens (red) of the phosphonate group (phosphorus-yellow). Tyr^{H33} moves 2.2 Å toward the phosphonate group, which is a key binding determinant in the hapten and is located in approximately the same position in the combining sites of the germline and affinity-matured Fab-hapten complexes.

bility or expression of the antibody (*24*).

Structural insights into the immune response. The structural analysis of 48G7 and its germline precursor suggests that the binding potential of the clonal antibody population may be significantly expanded as a result of the ability of a single germline antibody sequence to adopt more than one configuration (*4*, *26*). The crystal structures reveal differences between the structures of the germline Fab-hapten complex and the unbound Fab fragment that appear to be associated with hapten binding and not crystal packing. Such changes are not observed in the mature antibody, which involves a lock-and-key fit of hapten to the active site. The electron density of several neighboring residues also becomes more ordered on binding of hapten to the germline antibody (in contrast, the liganded and unliganded forms of the mature antibody are highly ordered). Thus, the binding of hapten 1 to the germline antibody results in a combining-site configuration with enhanced complementarity to the hapten as suggested by the "chemical instruction" model proposed by Pauling (*4*). Somatic mutation rather than folding of the remainder of the antibody molecule appears to stabilize this active site configuration. The ability of an antibody active site to reconfigure in response to antigen binding may significantly expand the structural diversity of the primary repertoire beyond that calculated from a consideration of sequence diversity alone. Previous kinetic studies have also provided evidence that a small subset of antibodies from the secondary and tertiary responses to 2-phenyl-5-oxazolone can adopt more than one conformation in the absence of ligand (*26*). It is therefore likely that both sequence and configurational diversity contribute to the ability of the immunoglobulin fold to bind an almost infinite array of chemical structures.

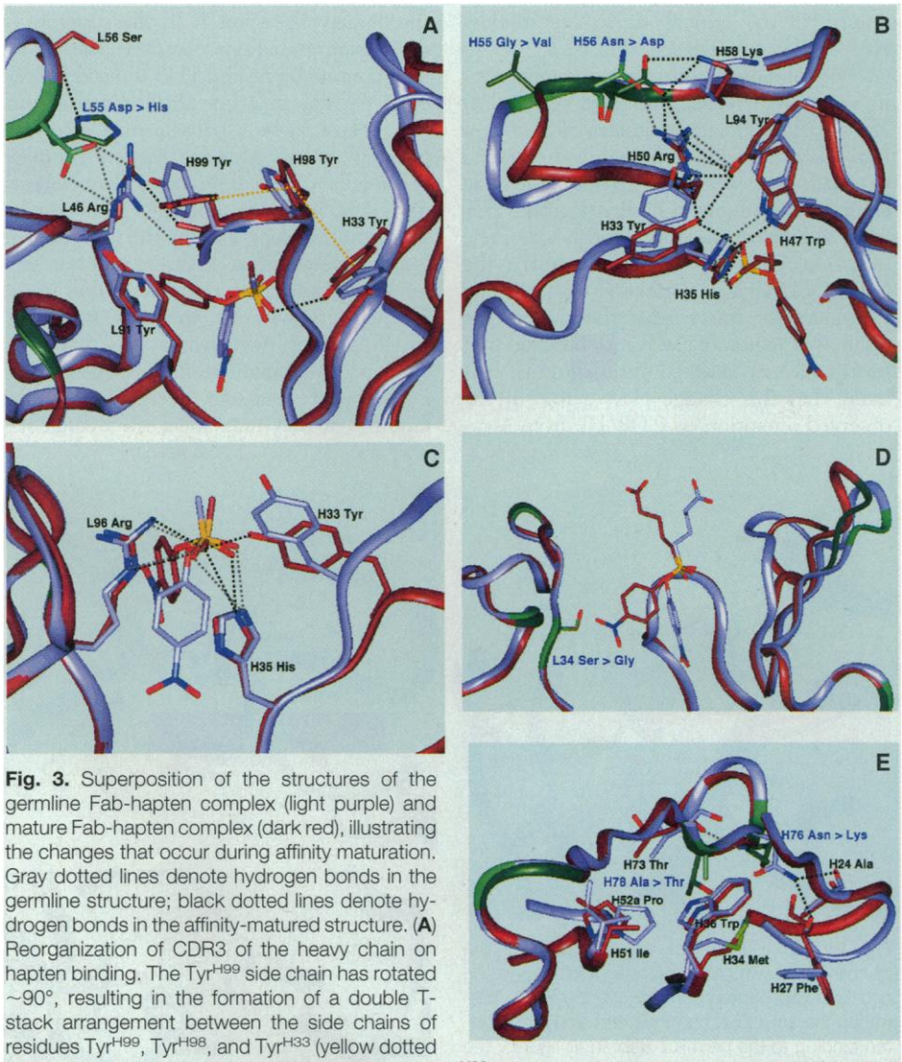


Fig. 3. Superposition of the structures of the germline Fab-hapten complex (light purple) and mature Fab-hapten complex (dark red), illustrating the changes that occur during affinity maturation. Gray dotted lines denote hydrogen bonds in the germline structure; black dotted lines denote hydrogen bonds in the affinity-matured structure. (**A**) Reorganization of CDR3 of the heavy chain on hapten binding. The Tyr^{H99} side chain has rotated ~90°, resulting in the formation of a double T-stack arrangement between the side chains of residues Tyr^{H99}, Tyr^{H98}, and Tyr^{H33} (yellow dotted lines); the additional hydrogen bond between the Tyr^{H33} hydroxyl and phosphonyl oxygen of the hapten is also shown. The mutation of Asp^{L55} → His has abolished hydrogen bonding interactions between that residue and Arg^{L46}; nonetheless, a hydrogen bond between the Arg^{L46} guanidinium group and the main-chain carbonyl of Tyr^{H99} remains and a hydrogen bond between His^{L55} and the backbone carbonyl of Ser^{L56} is formed. (**B**) The two somatic mutations, Gly^{H55} → Val and Asn^{H56} → Asp (shown in green), reorganize CDR1, CDR2 and CDR3 of the heavy chain in a series of interactions that influence residues in direct contact with the hapten molecule. The ε-amino group of Lys^{H58} forms a salt bridge to the Asp^{H56} carboxylate group, which in turn interacts with the Arg^{H50} guanidinium group. Additional hydrogen bonds between Arg^{H50}, Tyr^{H94} and Tyr^{H33} stabilize the orientation of the Tyr^{H33} side chain. The hapten has been omitted from the germline structure for clarity. (**C**) In the structure of the affinity-matured Fab-hapten complex, the Tyr^{H33} side chain is able to form an additional hydrogen bond to the phosphonate group. The phosphonate group and side chains of Arg^{L96} and His^{H35} are closer in the structure of the mature versus germline Fab. Note also that the phosphorus atom (yellow) in both structures occupies the same position in space relative to the surrounding combining-site residues, but that the nitrophenyl group (and aliphatic linker, not shown) are in different orientations. (**D**) The side chain of Ser^{L34} (somatic mutation Ser^{L34} → Gly) would interfere with the binding orientation of the hapten in the mature Fab-hapten complex (the side chain hydroxyl group of Ser^{L34} is 2.1 Å from the nitrophenyl oxygen). (**E**) The somatic mutations, Asn^{H76} → Lys and Ala^{H78} → Thr (green), control the location of CDR1 of the heavy chain by hydrogen-bonding interactions between Asn^{H76} and the main-chain carbonyls of Ala^{H24} and Phe^{H27} (germline), packing interactions between Thr^{H78}, Met^{H34}, Trp^{H36}, and Ile^{H51} (matured), and a hydrogen bond between the side chain of Lys^{H76} and the backbone carbonyl of Thr^{H73} (matured).

Somatic mutation of residues in the variable region also increases diversity and leads to enhanced affinity in clonally expanding B cell populations (*5*–*7*). Mutation at the active site can increase affinity for hapten. In most structurally characterized antigen-antibody complexes (*20*), either one or both of the CDR3 loops contacts antigen, as expected on the basis of their location at the center of the antibody combining site. Because of the combinatorial and nontemplated nature of the mechanisms that generate CDR3 (*6*), this central region of the antibody combining site is far more diverse than the flanking germline-encoded CDR1 and CDR2 regions in the primary antibody repertoire (*8*), paralleling the distribution of diversity observed in the T cell receptor (*27*). Unlike T cell receptors, antibodies depend on further diversification to drive affinity maturation, a process that is essential for generating neutralizing antibodies. The crystal structures of 48G7 and its germline precursor show how somatic mutations throughout the entire variable region can further optimize and stabilize the combining-site configuration induced by hapten.

Many of the somatic mutations reconfigure active site residues involved in binding interactions with hapten by reorganizing net-

works of hydrogen bonding, electrostatic, and van der Waals interactions between variable region residues over distances of 15 Å. This reorganization involves changes in both amino acid side chain interactions and backbone conformation [which have also been observed in somatically related anti-*p*-azophenylarsonate Fabs (9)]. This process may be facilitated by the particular architecture of the variable region, such that the packing of loops (8) against one another makes possible many alternative networks of side chain interactions. Such interactions may not be so easily propagated throughout a variable region consisting of α helices or β sheets. The end result of these somatic mutations is a combining site with improved complementarity to hapten (including an additional hydrogen bond to the key phosphonyl group of hapten) which, in contrast to the germline antibody, binds hapten in a pre-organized fashion. The latter suggests that, in addition to enthalpic effects, entropic restriction of residues in the combining site has a key role in the 30,000-fold increase in binding affinity which occurs during affinity maturation. The crystal structure data further support the view (*10*) that the improvement in affinity is the result of many small additive changes rather than a few large effects. This observation underscores the importance of multivalent display of antigen on follicular dendritic cells in germinal centers, the site of affinity maturation (*28*). This architecture allows for an amplification of small improvements in binding affinity to be transduced into large changes in signals required for cell survival and proliferation in germinal centers.

The catalytic antibody 48G7 represents the only example in which the binding properties of both a germline and affinity matured antibody have been investigated at a detailed structural level. However, this analysis may help to explain the fundamental issue of how the immune system copes with an unlimited number of antigens. One can speculate that in addition to the clonal nature of the immune response, many germline antibodies may indeed adopt multiple configurations with antigen binding, together with somatic mutation, stabilizing the configuration with optimum complementarity to antigen. The degree to which an individual germline antibody exploits the configurational diversity described above will likely depend on the initial fit of antigen to germline antibody and the nature of the forces driving antigen-antibody complexation. Similar analyses of other antibodies with diverse binding properties, including those that exhibit polyspecificity (*29*) will provide further insights into the molecular mechanisms of the immune response.

The capacity of antibodies to bind antigens and, with the proper chemical instruction, also catalyze chemical reactions (*1*) suggests that the same principles described above may have played an important role in the early evolution of enzymes. One can envisage that a relatively limited number of protein frameworks, each with the ability to adopt many different active or combining site geometries in response to both ligand binding (*24*) as well as mutations throughout the protein structure, may have provided an efficient means of evolving a diverse range of substrate specificities and catalytic functions. Finally, if we consider the immune system as a paradigm for other combinatorial approaches to evolving new function, the lessons derived from this study may prove useful in designing new strategies for generating and presenting chemical diversity (*30*).

REFERENCES AND NOTES

1. S. J. Pollack, J. W. Jacobs, P. G. Schultz, *Science* **234**, 1570 (1986); A. Tramontano, K. D. Janda, R. A. Lerner, *ibid.*, p. 1566; P. G. Schultz and R. A. Lerner, *ibid.* **269**, 1835 (1995); ———, *Acc. Chem. Res.* **26**, 391 (1993).
2. See L. C. Hsieh-Wilson, X.-D. Xiang, P. G. Schultz, *Acc. Chem. Res.* **29**, 391 (1996).
3. F. Breinl and F. Haurowitz, *Zs. Physiol. Chem.* **192**, 45 (1930); F. Haurowitz, *Cold Spring Harbor Symp. Quant. Biol.* 559 (1967).
4. L. Pauling, *J. Am. Chem. Soc.* **62**, 2643 (1940).
5. F. M. Burnet, *The Clonal Selection Theory of Acquired Immunity* (Vanderbilt Univ. Press, Nashville, TN, 1959) p. 53; D. W. Talmage, *Science* **129**, 1649 (1959); J. Lederberg, *ibid.*, p. 1649 (1959).
6. S. Tonegawa, *Nature* **302**, 575 (1983).
7. D. L. French, R. Laskov, M. D. Scharff, *Science* **244**, 1152 (1989).
8. C. Chothia *et al.*, *J. Mol. Biol.* **227**, 799 (1992); I. M. Tomlinson *et al.*, *ibid.* **256**, 813 (1996).
9. R. K. Strong, G. A. Petsko, J. Sharon, M. N. Margolies, *Biochemistry* **30**, 3749, (1991).
10. P. A. Patten *et al.*, *Science* **271**, 1086 (1996); G. J. Wedemayer, L. H. Wang, P. A. Patten, P. G. Schultz, R. C. Stevens, *J. Mol. Biol.* **268**, 390 (1997).
11. R. C. Stevens *et al.*, *Israel J. Chem.* **36**, 121 (1996).
12. L. C. Hsieh-Wilson, P. G. Schultz, R. C. Stevens, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5363 (1996).
13. H. Ulrich *et al.*, in preparation.
14. J. W. Jacobs, thesis, University of California, Berkeley (1989).
15. Crystals of the germline Fab in the presence and absence of hapten were obtained by the hanging drop method (*31*). Protein solutions of Fab consisted of 10 to 15 mg/ml in 100 mM NaCl, 10 mM tris, pH 8.0, 1 mM methionine, 1 mM sodium azide, and 0.5 mM EDTA. The mother liquors were 2.0 M $(NH_4)_2SO_4$ for germline Fab crystals without hapten, and 0.1 M ammonium acetate, 0.1 M Na cacodylate, pH 6.0, 18 percent PEG 4000, 1 percent dioxane for the germline Fab in the presence of hapten. In the latter case, approximately 50 molar excess hapten was added to the protein solution to a final concentration of 10 mM. The complex crystallized in the space group $P2_1$ with unit cell parameters a = 46.1 Å, b = 60.9 Å, c = 73.1 Å, β = 104.8°. The Fab without hapten crystallized in the space group $P2_12_12_1$ with unit cell parameters a = 66.2 Å, b = 77.6 Å, c = 86.6 Å. For data collection, crystals were soaked for ~5 minutes in 20 percent glycerol, 80 percent mother liquor, and then frozen by plunging into liquid nitrogen. Data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) to a maximum resolution of 2.1 Å. The data were integrated and merged with the DENZO/SCALEPACK package (*32*). The structure was determined by molecular replacement using a modified mature 48G7 with hapten structure (*10*). The nine somatic point mutations were replaced as alanines and used as a search model. The rotation and translation search was performed with AMORE (*33*). Both germline data sets provided a strong peak solution and the best solution was placed into the program X-PLOR (*34*) for rigid-body minimization. The model was completed with several cycles of molecular dynamics by X-PLOR and subsequent model building with program O (*35*). Free *R* values were monitored throughout the model-building. A bulk solvent correction was applied with the X-PLOR. All CDR loops were omitted and rebuilt into density, and the hapten was fitted into the germline with the hapten data set (Table 1). Model quality was checked using the program PROCHECK (36).
16. The germline precursor and mature 48G7 used for the biophysical, kinetic, and crystallographic studies were expressed as a chimeric Fab in which the V_H and V_L variable region genes are fused to the human C_H1 and C_κ constant regions, respectively (*37*).
17. The V_L-V_H rotation angle was calculated [R. L. Stanfield, M. Kamimura, J. M. Rini, A. T. Profy, I. A. Wilson *Structure* **1**, 83 (1993)]. The heavy chains of two Fab variable domains were superimposed by the program XPLOR (*34*). The rotation angle was defined as the rotation angle required to bring the light chains into coincidence.
18. E. Fischer *Ber. Dtsch. Chem. Ges.* **27**, 2985 (1984).
19. S. K. Burley and G. A. Petsko, *Science* **229**, 23 (1985); D. A. Dougherty and D. A. Stauffer, *ibid.* **250**, 1558 (1990).
20. I. A. Wilson and R. L. Stanfield *Curr. Opinion Struct. Biol.* **4**, 857 (1994); R. A. Mariuzza and R. J. Pollack, *ibid.* **5**, 50 (1993); D. R. Davis and S. Chacko *Acct. Chem. Res.* **26**, 421 (1993); D. M. Webster, A. H. Henry, A. R. Rees, *Curr. Opinion Struct. Biol.* **4**, 123 (1994).
21. C. Chothia *et al.*, *Nature* **342**, 877 (1989).
22. C. M. Wilmot and J. M. Thornton, *J. Mol. Biol.* **203**, 221 (1988).
23. G. N. Ramachandran and V. Sasisekharan, *Adv. Protein Chem.* **23**, 283 (1968).
24. M. Dueñas *et al.*, *Gene* **158**, 61 (1995); M. R. Hurle, L. R. Helms, L. Li, W. Chan, R. Wetzel, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5446 (1994); H. Yasui, W. Ito, Y. Kurosawa, *FEBS Lett.* **353**, 143 (1994); K. Alfthan *et al.*, *Gene* **128**, 203 (1993).
25. D. E. Koshland Jr., *Angew. Chem. Int. Ed. Engl.* **33**, 2375 (1994).
26. J. Foote and C. Milstein, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10370 (1994); D. R. Davies, E. A. Padlan, S. Sheriff *Annu. Rev. Biochem.* **59**, 439 (1990).
27. M. Davis and P. Bjorkman, *Nature* **334**, 395 (1988).
28. L. MacLenran, *Annu. Rev. Immunol.* **12**, 117 (1994).
29. M. Fougereau and C. Schiff, *Immunol. Rev.* **105**, 69 (1988); A. B. Hartman *et al.*, *Mol. Immunol.* **26**, 359 (1989); E. A. Kabat and T. T. Wu, *J. Immunol.* **147**, 1709, (1991); S. Ghosh and A. M. Campbell, *Immunol. Today* **7**, 217 (1986). F. Romesberg *et al.*, unpublished results.
30. A. Cramier, E. Whitehorn, E. Tate, W. Stemmer, *Nature Biotechnol.* **14**, 315 (1996).
31. J. Jancarik and S.-H. Kim, *Appl. Cryst.* **24**, 409 (1991).
32. Z. Otwinowski and W. Minor, in preparation.
33. J. Navaza, *Acta Cryst. A* **50**, 157 (1994).
34. A. T. Brünger, *X-PLOR 3.8* (Yale Univ. Press, New Haven, CT, 1996).
35. T. A. Jones and K. Kjeldgaard, "O" Computer Graphics Program (Uppsala Univ., Sweden, 1993).
36. R. A. Laskowski *et al.*, *J. Appl. Cryst.* **26**, 282 (1993).
37. P. Carter *et al.*, *Bio/Technology* **10**, 163 (1992); H. Ulrich *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11907 (1995).

38. We thank I. Wilson, B. Spiller, and B. Santarsiero for helpful comments on the manuscript. Supported by DOE contract DEACO3765F00098 (P.G.S.), NIH grant R01 AI39089 (R.C.S.), and the Howard Hughes Medical Institute (P.G.S.). We thank the Stanford Synchrotron Radiation Laboratory for Synchrotran beamtime. The coordinates have been deposited in the protein data bank with accession numbers 1AJ7 and 2RCS.

12 November 1996; accepted 18 March 1997