The Immunological Evolution of Catalysis

Phillip A. Patten, Nathanael S. Gray, Priscilla L. Yang, Cara B. Marks, Gary J. Wedemayer, J. Jay Boniface, Raymond C. Stevens,* Peter G. Schultz*

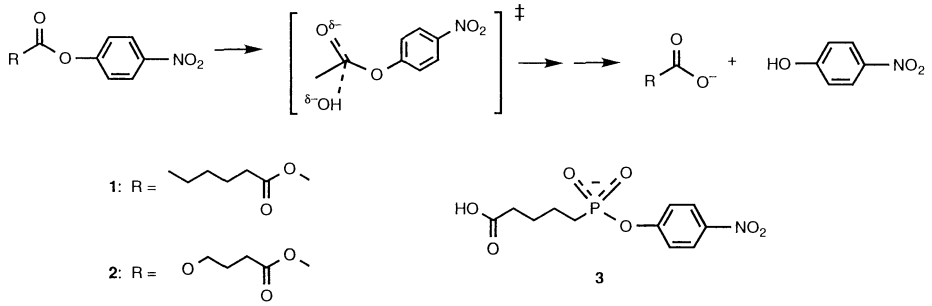
The germline genes used by the mouse to generate the esterolytic antibody 48G7 were cloned and expressed in an effort to increase our understanding of the detailed molecular mechanisms by which the immune system evolves catalytic function. The nine replacement mutations that were fixed during affinity maturation increased affinity for the transition state analogue by a factor of 10^4, primarily the result of a decrease in the dissociation rate of the hapten-antibody complex. There was a corresponding increase in the rate of reaction of antibody with substrate, k_{cat}/K_m, from 1.7×10^2 M^{-1} min^{-1} to 1.4×10^4 M^{-1} min^{-1}. The three-dimensional crystal structure of the 48G7-transition state analogue complex at 2.0 angstroms resolution indicates that none of the nine residues in which somatic mutations have been fixed directly contact the hapten. Thus, in the case of 48G7, affinity maturation appears to play a conformational role, either in reorganizing the active site geometry or limiting side-chain and backbone flexibility of the germline antibody. The crystal structure and analysis of somatic and directed active site mutants underscore the role of transition state stabilization in the evolution of this catalytic antibody.

The immune response provides a means to rapidly generate high affinity, selective receptors for an enormous array of chemical structures. The primary immune response occurs through the generation and screening of a large and diverse library of antibody combining sites, resulting from the combinatorial rearrangement of variable (V), diversity (D), and joining (J) gene segments (*1*). Antigen-antibody recognition is optimized during the secondary and tertiary responses through the somatic hypermutation of antibody genes (*2*) in clonally expanding populations of B cells, with concomitant selection between competing clones for increasingly rare antigen (*3*). This process results in a population of rapidly evolving B cells that produce antibodies of increasing affinity as the immune response proceeds, giving rise to the well-known phenomenon of affinity maturation (*4*, *5*).

There are parallels between this process and the natural evolution of enzyme active sites. The latter also arise by a process involving the generation of molecular diversity, in this case by exon shuffling and point mutations (*6*), coupled with a selection generally based on catalytic efficiency. Whereas antibodies are selected on the basis of affinity for stable antigens, enzymes evolve based on catalytic efficiency which in turn is correlated with affinity for high energy, transition states (*7*, *8*). The similarities between natural selection and the immune response suggest that, with proper chemical instruction, the latter can be directed along the same pathway as enzymatic evolution to afford catalytic antibodies. Indeed, when stable transition state analogues are used as immunogens, antibodies can be generated that catalyze a large number of different reactions, from disfavored cyclization reactions to pericyclic and redox reactions (*9*, *10*). In some cases, the rates of the antibody-catalyzed reactions approach those of comparable enzyme-catalyzed reactions (*11*), and in others, antibodies have been generated that catalyze reactions for which there is no known enzymic counterpart (*12*).

In order to reconstruct the immunological evolution of a catalytic antibody, we cloned the germline genes of the esterolytic antibody 48G7. Expression of the corresponding antibody (*13*, *14*) allowed us to reconstruct the immunological evolution of this catalytic antibody by characterizing the functional consequences of affinity maturation on hapten binding and catalysis. In addition, the three-dimensional x-ray crystal structure of the Fab fragment of 48G7, complexed with the transition state analogue against which it was elicited, was solved at 2.0 Å resolution. This structure, along with a mutagenesis study of active site residues, has provided insight into the mechanism of this antibody-catalyzed reaction. The crystal structure has also made it possible to begin to interpret the effects of somatic mutation in structural and functional terms.

Functional characterization of the germline antibody. Antibody 48G7, which binds the *p*-nitrophenyl phosphonate transition state analogue **3**, catalyzes the hydrolysis of the corresponding *p*-nitrophenyl ester **1** and carbonate **2** with rate accelerations of 1.6×10^4 and 4×10^4, respectively, compared to the rates of uncatalyzed reaction (*14*). The antibody-catalyzed reaction is first order in hydroxide ion from pH 6.2 to 9.2 for substrate **2**, and chemical modification studies suggested that arginine, tyrosine, and histidine residues are important in catalysis (*14*). In addition, 48G7 has been cloned (*13*) and expressed in bacteria by fusing the V_H (heavy) and V_L (light) variable region genes to human C_H1 and C_κ constant regions, respectively (*15*, *16*). The values of the catalytic constant, k_{cat}, and Michaelis constant, K_m, of the purified chimeric Fab fragment for hydrolysis of ester **1** are 5.5 min^{-1} and 391 μM, respectively, comparable to those for the hybridoma-derived antibody (k_{cat} = 2.1 min^{-1}; K_m = 113 μM) (*14*).



Scheme 1. Antibody-catalyzed hydrolysis reaction and corresponding transition state analogue **3**.

To determine the degree to which the binding and catalytic properties of 48G7 preexisted in the germline antibody, or evolved as a consequence of affinity maturation, the germline light (L) and heavy chain (H) variable region genes were cloned and sequenced. Nine replacement mutations, three in the light chain and six

P. A. Patten, N. S. Gray, P. L. Yang, and P. G. Schultz are in the Howard Hughes Medical Institute, Department of Chemistry, University of California, Berkeley, CA 94720, USA. C. B. Marks, G. J. Wedemayer, and R. C. Stevens are in the Department of Chemistry, University of California, Berkeley, CA 94720, USA. J. J. Boniface is in the Howard Hughes Medical Institute, Department of Microbiology and Immunology, Stanford University, Stanford, CA 94305, USA.

*To whom correspondence should be addressed.

in the heavy chain, were fixed during the affinity maturation process (Fig. 1). The germline variable region genes were then combined with their respective D and J region sequences; the resulting Fab fragment was expressed in *Escherichia coli*, and its binding and catalytic properties were characterized (*15*, *16*). In addition, each of the reconstructed heavy and light chain germline genes (G^H and G^L) were expressed in combination with the partner from 48G7 to yield G^H48G7^L and G^L48G7^H.

A functional analysis of these somatic mutations revealed that both hapten binding and catalytic activity increased during affinity maturation. The dissociation constant (K_d) of the germline Fab · **3** complex is 135 μM compared to that of 10 nM for the 48G7 Fab · **3** complex, representing a 1.4×10^4 improvement in binding affinity during the affinity maturation process (Table 1). The half-germline Fabs G^L48G7^H and $48G7^LG^H$ have intermediate K_d's of 330 nM and 1860 nM, respectively. Paralleling this improvement in affinity, the k_{cat}/K_m value for substrate **1** (the bimolecular rate constant for reaction of free antibody with substrate) increased from 1.7×10^2 M^{-1} min^{-1} for germline Fab to 1.4×10^4 M^{-1} min^{-1} for 48G7. The low level of germline catalytic activity precluded measurement of k_{cat} and K_m independently. The increase in k_{cat}/K_m with increasing affinity for the transition state analogue is consistent with the concepts of enzymatic catalysis put forth by Pauling (*7*) and Haldane (*8*) in which preferential binding energy for the transition state lowers the activation energy for reaction of enzyme with substrate. Other factors (for example, binding of substrate and product and the environment around the attacking water molecule) that are less subject to selective pressure during the immune response may account for the imperfect correlation between binding affinity and catalytic efficiency. In fact, affinity maturation does lead to a moderate increase in affinity for the reaction product [K_i = 35 μM (*14*)] relative to the germline Fab (K_d > 200 μM). This leads to increased product inhibition and is reflected in nonlinear kinetics in the case of 48G7, but not in the germline or half germline Fabs nor in 48G7 at pH 7.0 (below the pK_a of nitrophenol). This result again emphasizes the importance of minimizing the similarity of the product to the designed hapten in order to maximize catalytic antibody rates.

The catalytic and binding properties of the half-germline Fab constructs are intermediate between the germline Fab and 48G7, indicating that mutations in both chains contribute significantly to the affinity maturation process. Indeed, the high degree of additivity of the binding free energies of the heavy chain mutations (G^L48G7^H) and light chain mutations ($48G7^LG^H$) for hapten show that in this system the light and heavy chains can be optimized independently during affinity maturation (Table 1). These results are similar to those obtained on the effects of somatic mutations, studied individually and in combination, on the affinity of a secondary phenylarsonate specific antibody (*17*). This observed additivity of binding free energies may help to explain how B cells can find binding optima during affinity maturation through an evolutionary process that very sparsely samples sequence space. There are more than 10^{16} ways to distribute nine single-step replacement mutations just in the complementarity determining regions (CDR's) of 48G7 and $>10^{22}$ nine single-step replacement mutations, if distributed

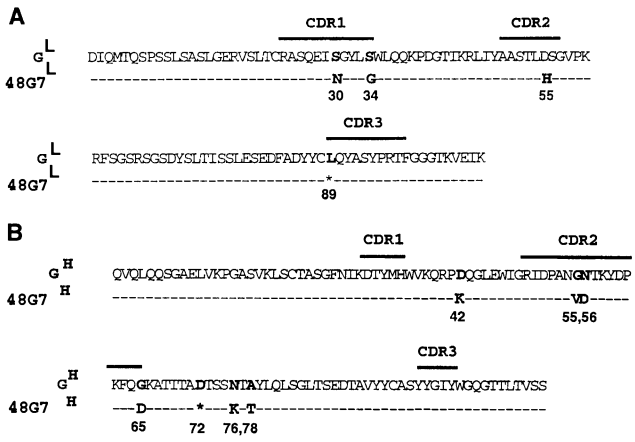


Fig. 1. Somatic mutations fixed during affinity maturation. The protein sequences of the germline (G^L, G^H) and 48G7 ($48G7^L$, $48G7^H$) variable domains of (**A**) the light and (**B**) the heavy chain are given. Dashes indicate identity with the germline sequence (silent mutations indicated with bold stars). Residue numbering and CDR's (bars) are as defined by Kabat (*41*). The nucleotide differences between the $48G7^L$ (*13*) and G^L genes at positions 30, 34, 55, and 89 (silent) are: AGT → AAT, AGC → GGC, GAT → CAT, and CTA → CTG, respectively (*42*). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Table 1. Catalytic and binding properties of 48G7 mutants. The catalytic assays and BIAcore analysis were performed as described (*15*, *16*, *43*, *44*). ND, not determined.

Fab construct		k_{cat}	K_m	k_{cat}/K_m
Light chain	Heavy chain	(min^{-1})	(μM)	(min^{-1} M^{-1})
		Active site mutants		
48G7	48G7	5.5	391	1.4×10^4
R96Q	48G7	0.5	140	3.6×10^3
R96K	48G7	1.1	220	5.0×10^3
48G7	Y33F	1.7	600	2.8×10^3
48G7	H35E	0.2	830	2.4×10^2
48G7	H35Q	3.3	700	4.7×10^3

Fab construct		k_{cat}	K_m	k_{cat}/K_m	K_d (hapten)	k_{on}	k_{off}	$\Delta G°_{TSA}$	$\Delta\Delta G°_{TSA}$
Light chain	Heavy chain	(min^{-1})	(μM)	(min^{-1} M^{-1})	(μM)	(M^{-1} sec^{-1})	(sec^{-1})	(kcal/mol)	(kcal/mol)
48G7	48G7	5.5	391	1.4×10^4	0.0045 (.010)*	1.45×10^5	1.41×10^{-3}	10.9	5.6
Germline	48G7	<1†	†	6.0×10^2	0.33	7.5×10^5‡	0.246	8.8	3.5
48G7	Germline	0.83	634	1.3×10^3	1.86	3.2×10^5‡	0.606	7.8	2.5
Germline	Germline	<1†	†	1.7×10^2	135	ND	>1	5.3	0

*The K_d of 0.010 μM is derived from the ratio of the rate constants k_{on} and k_{off}. A K_d of 0.010 ± 0.001 μM was obtained on three independent runs using three independently derivatized BIAcore chips. The k_{on} was determined at five different Fab concentrations. The ΔG's and $\Delta\Delta G$'s relative to germline were calculated from the K_d's. †Fab's for which accurate deconvolution of k_{cat} and K_m was not possible. ‡The G^L48G7^H and $48G7^LG^H$ k_{on} rate constants were calculated from the equation $k_{off} = k_{on}/K_d$ (*44*). The values of k_{cat}/K_m were derived from the initial slope of V as a function of the concentration of *S* [S] and in each case χ^2 >0.99. The kinetic parameters, k_{cat} and K_m, were determined by fitting of the data to the Michaelis Menten equation with nonlinear regression obtained with the program Kaleida Graph.

over the entire variable region (*18*). If most favorable somatic mutations are additive and insensitive to changes at other sites, then affinity maturation consists of an adaptive walk on a "correlated" fitness landscape (*19*), and a million B cells per clone sampling approximately one somatic mutation per cell division can result in significant enhancements in affinity (*20*).

In order to assess whether improvements in hapten binding during the affinity maturation process are correlated with changes in either the rate constants for Fab-hapten complex formation (k_{on}), dissociation (k_{off}), or both, the binding of Fab to hapten was studied by surface plasmon resonance techniques. The results of this analysis (Table 1) show that the increase in affinity for hapten that arises during affinity maturation results primarily from a decrease in the apparent k_{off} from >1 sec^{-1} for the germline Fab fragment to 1.4×10^{-3} s^{-1} for the 48G7 Fab fragment. Apparent k_{on} values were very similar and range from 1.45×10^5 M^{-1} s^{-1} to 7.5×10^5 M^{-1} s^{-1}, which are typical of hapten-antibody association rates (*21*, *22*).

Structural analysis. The structure of the recombinant 48G7 Fab fragment was determined at 2.0 Å resolution by molecular replacement (*23*). The designate of the hapten was observed with an F_o-F_c omit map, with the electron density of the phosphonate group being particularly strong. The final model for the structure is listed in Table 2. The structure of 48G7 is similar to those of other known antibodies. The three light chain CDRs and CDR1 of the heavy chain in the crystal structure all adopt the predicted canonical structures for these loops, based on the primary structure of the antibody (*24*, *25*). The primary structure of the germline heavy chain CDR2 conforms to canonical structure 2 but the hydrogen-bonding pattern is altered due to somatic mutation of Gly^{55} at the $i+4$ position of the loop to valine. Heavy chain CDR3 of 48G7 is relatively short and forms a β turn between residues 98 and 101.

The hapten is bound in a pocket roughly 10 Å deep (Fig. 2) with the aromatic nitro group at the bottom of the pocket and the aliphatic linker and carboxylic acid group at the surface surrounded by six tyrosine residues (Tyr^{32L}, Tyr^{91L}, Tyr^{94L}, Tyr^{33H}, Tyr^{99H}, Tyr^{100H}), similar to the "aromatic cluster" noted in an antibody to phenylarsonate (*26*). Approximately 80 percent of the surface area of the hapten is buried in the Fab-hapten complex. The primary hapten-antibody interactions are between the hapten and CDR3 of the light chain and CDR1 and CDR3 of the heavy chain (Fig. 2). The aliphatic linker of the hapten makes van der Waals contact with Tyr^{99H} and Tyr^{100H}, and the nitrophenyl ring is in contact with Tyr^{91L}. No residues directly contact the aromatic nitro group. The anionic phosphonate group is bound by three hydrogen-bonding interactions: a salt bridge between Arg^{96L} and the pro *S*-phosphonyl oxygen of **3**, a hydrogen bond between Tyr^{33H} and the same phosphonyl oxygen, and a hydrogen bond between the ε-imino group of His^{35H} and the pro *R*-phosphonyl oxygen of **3** (Fig. 2). The importance of His^{35H}, Tyr^{33H}, and Arg^{96L} in hapten binding is in agreement with earlier chemical modification experiments on 48G7 (*14*). These residues are oriented toward the phosphonate moiety via a series of secondary hydrogen-bonding interactions involving Arg^{96L} and the Tyr^{91L} backbone carbonyl group, Tyr^{33H} and the Tyr^{94L} hydroxyl group, and His^{35H} and the imino group of the conserved Trp^{47H}.

The active site structure of the Fab-hapten complex does not reveal any residue that is an obvious candidate for an active site nucleophile, consistent with stopped-flow kinetic studies for which there was no evidence of an initial burst of nitrophenolate ion. Thus, direct attack by an active site water or hydroxide anion probably leads to hydrolysis. A likely mechanism, based on the crystal structure, is that His^{35H}, Tyr^{33H}, and Arg^{96L} stabilize the negatively charged tetrahedral transition state via a combination of electrostatic and hydrogen-bonding interactions. Alternatively, His^{35H} may act

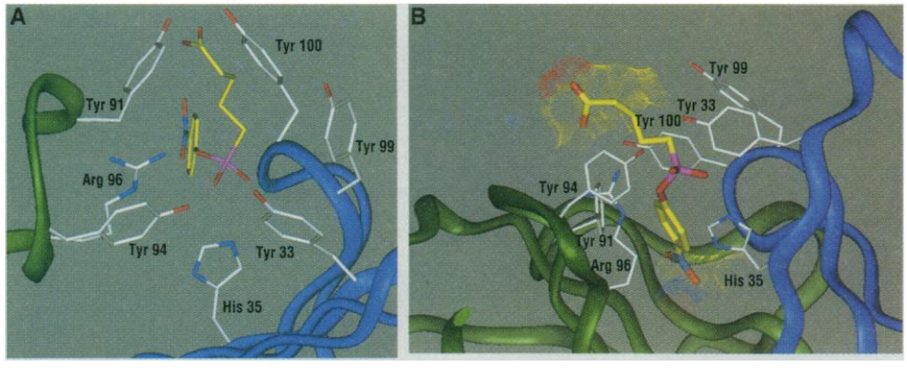


Fig. 2. Close-up view of the 48G7 active site with the nitrophenyl phosphonate transition state analogue bound. (**A**) The active site residues in the light chain (green) (Tyr^{91L}, Tyr^{94L}, Arg^{96L}) and the heavy chain (blue) (Tyr^{33H}, His^{35H}, Tyr^{99H}, Tyr^{100H}) are shown (atom colors: O, red; N, blue; C, white). The hapten molecule has the carbon atoms shown in yellow to distinguish it from the rest of the molecule. (**B**) Alternative view of the hapten active site with a Connelly surface of the solvent-accessible region of the hapten molecule displayed. The majority of solvent accessibility of the hapten is in the aliphatic linker, with a small percentage in the nitro group of the nitrophenyl moiety.

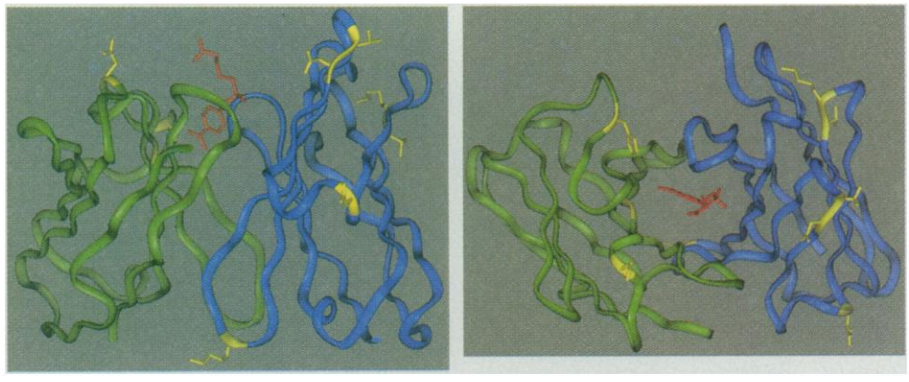


Fig. 3. Location of somatic mutations in the 48G7 structure. Two ribbon diagram views of the 48G7 Fab fragment are shown with the light chain in green, the heavy chain in blue, and the bound hapten in red. The side chains of the somatic mutations are indicated in yellow (Ser^{30L} → Asn, Ser^{34L} → Gly, Asp^{55L} → His, Glu^{42H} → Lys, Gly^{55H} → Val, Asn^{56H} → Asp, Gly^{65H} → Asp, Asn^{76H} → Lys, Ala^{78H} → Thr). The closest contact between the hapten and a somatic mutation is with His^{55L} (green chain), 5.3 Å away.

Table 2. Data collection and refinement statistics.

Item	48G7 with hapten
Number of observations	121,262
Number of unique reflections	33,282
Completeness to refined resolution (%) (1.95)	96.7
R_{sym} (*I*)(%)	7.2
Average I/σ	8.1
Refinement resolution	6.0 to 1.95
R_{cryst} (%)	24.3
rms Δ bond lengths (Å)	0.007
rms Δ bond angles (degrees)	1.613

as a general base to activate a water molecule for attack on the acyl group of the complexed substrate. It is not possible to discern an inflection in the pH dependence of the antibody-catalyzed reaction corresponding to protonation of His^{35H} because of the low rate of the reaction in this pH range. The absence of a proton donor near the leaving nitrophenolate anion may explain the inability of 48G7 to catalyze the hydrolysis of the corresponding nitroanilide or carbamate substrates (*14*, *27*).

A number of active site mutants were generated and characterized (Table 1). Mutation of Arg^{96L} to glutamine led to a drop in k_{cat} from 5.5 min^{-1} to 0.5 min^{-1}, and mutation of Tyr^{33H} to phenylalanine resulted in a reduction in k_{cat} to 1.7 min^{-1}. Although these results are consistent with the postulated role of these residues in stabilizing the negatively charged transition state, the magnitude of the effects is less than expected. Because both Arg^{96L} and Tyr^{33H} hydrogen bond to the pro S oxygen of the $TS^{‡}$ analogue, deletion of one hydrogen bond donor may be partially compensated by the second residue. Mutation of His^{35H} to glutamate led to reduction in k_{cat} by a factor of 30, indicating a significant catalytic or structural role for this residue. Mutation of His^{35H} to glutamine had relatively little effect on catalysis, suggesting that His^{35H} acts by stabilizing the negatively charged transition state. Further analysis of these and other mutants should provide more insight into the mechanism of catalysis and could lead to mutants with enhanced activity.

A comparison to other esterolytic antibodies. Two other crystal structures of esterolytic antibodies with bound phosphonate (*28*) or phosphate (*29*) transition state analogues have been solved. The active site of the antibody 17E8 forms an oxyanion hole in which positively charged residue Lys^{97H}, and possibly Arg^{96L}, stabilize the negatively charged transition state as in the case of 48G7. However, in antibody 17E8, an active site serine residue, Ser^{99H} (the analogous residue is a Tyr in 48G7), rather than a water molecule, may act as the nucleophile. In both 48G7 and 17E8, His^{35H} appears to play a significant mechanistic role. In McPC603 the negatively charged phosphate is stabilized by an oxyanion hole similar to that seen in 48G7, consisting of Tyr^{33H} and Arg^{52H} (*29*). A third structure of a phosphonate-specific esterolytic antibody has been solved at 3 Å resolution in the absence of hapten (*30*). Modeling of the $TS^{‡}$ analogue into the active site suggests a possible hydrogen-bonding interaction between two tyrosine hydroxyl groups, Tyr^{96L} and Tyr^{50H}, and the negatively charged phosphonate group. Again, His^{35H} is relatively close to the phosphonate group. Finally, a three-dimensional model has been constructed for the esterolytic antibody NPN43C9, which was elicited by a phosphonamidate transition state analogue (*31*). In this model, which is supported by mutagenesis data, Arg^{96L} forms a salt bridge with the tetrahedral negatively charged phosphonamidate group, analogous to the Arg^{96L}-phosphonate interaction seen in 48G7. Thus, although these catalytic antibodies were all generated independently with different transition state analogues, they all share the common feature of an oxyanion binding site, underscoring the importance of transition state stabilization in these esterolytic antibodies. The interaction of these negatively charged phosphorus-containing $TS^{‡}$ analogues with a positively charged $VJ_{κ}$ junctional residue in 48G7, 17E8, and NPN43C9 indicates a strong selection for this structural feature and resembles the occurrence of Arg^{96} ($V_{κ}$-$J_{κ}$ junction) in 48 of 48 phenylarsonate specific antibodies (*32*) as well as the strong selection for charge complementarity between bound peptide antigens and T cell receptor CDR3 loops (*33*). Variations in these antibody structures, which are reflected in their different catalytic properties (*28*, *30*, *31*), indicate the diversity of the immune response and provide additional opportunities to examine the effects of structural variations on catalytic function.

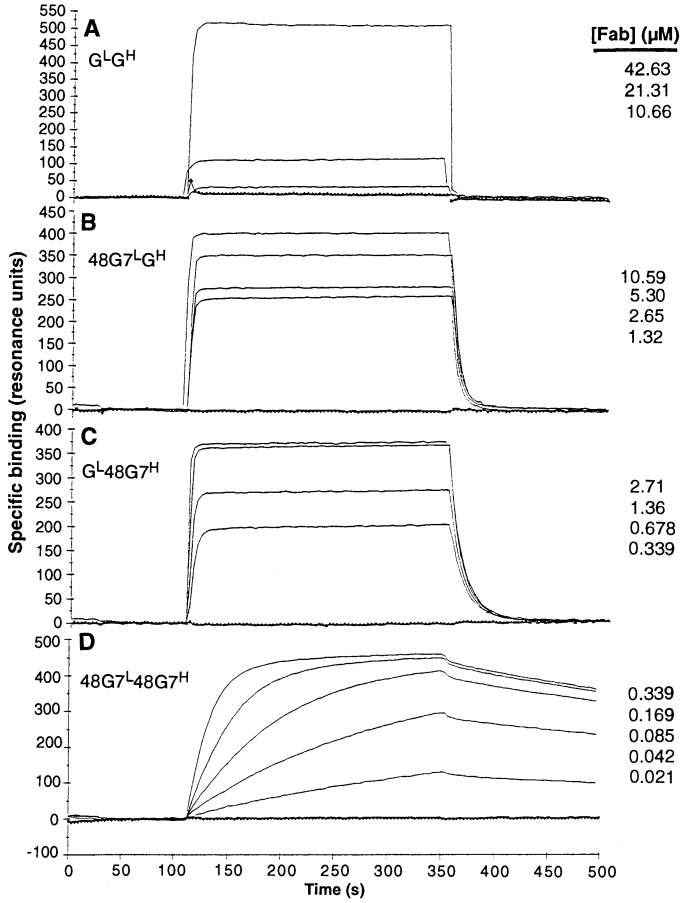


Fig. 4 (A to D). BIAcore analysis of affinity maturation. Binding kinetics of purified Fab's to immobilized **3** were determined as indicated (*43*). The concentrations of Fab are indicated in the column on the right. Bulk phase changes and nonspecific binding have been eliminated by subtracting out sensorgrams from negative control surfaces such that the sensorgrams indicate only hapten-specific binding. Data collected in the presence of 1 mM hapten **3** (triangles) indicate that all specific binding is blocked by free hapten.

Structural aspects of affinity maturation. Examination of the crystal structure of the 48G7 Fab-hapten complex indicates that none of the somatically mutated residues is in direct contact with bound hapten **3**. It is likely that a number of these mutations are involved in hapten binding in view of the large increase in the affinity for hapten **3**. This is supported by the properties of the half-germline antibodies as well as a point mutational analysis of the light chain somatic mutations in which the mutations G34S and H55D each resulted in the affinity for **3** becoming eight times lower.

These somatic mutations likely play an important role in organizing the conformation of residues that make up the active site of 48G7. For example, the $Ala^{78H} \rightarrow$ Thr and $Asn^{76H} \rightarrow$ Lys substitutions in the D to E loop pack against CDR1 and may modulate the geometries of the adjacent active site residues, Tyr^{33H} and His^{35H}. The substitutions $Gly^{55H} \rightarrow$ Val and $Asn^{56H} \rightarrow$ Asp lie at the base of a turn (Pro^{52H}-Ala^{53H}-Asn^{54H}-Val^{55H}) some 15 to 20 Å from the bound hapten. The change from glycine to the β-branched valine at position $i+4$ in the turn may affect the packing of this loop against CDR1 of the heavy chain, which contains the active site residues Y^{33H} and H^{35H}. In addition, Asp^{56H} is hydrogen-bonded to Arg^{50H} which in turn hydrogen

bonds to Tyr^{94L}, an active site residue responsible for fixing the oxyanion hole conformation. The somatically mutated residues, $Ser^{34L} \rightarrow Gly$ and $Asp^{55L} \rightarrow His$, lie at the bottom of the binding site for the nitrophenyl ring (although Gly^{34L} does not contact hapten, S^{34L} in the germline may sterically hinder hapten binding). The $Ser^{30L} \rightarrow Asn$ change is at the $i+1$ position of a solvent-exposed type I turn, the $Gly^{65H} \rightarrow Asp$ change lies in a turn distant from the active site and the $Glu^{42H} \rightarrow Lys$ change lies in a turn at the interface of the heavy and light chains distant from the active site. Although the crystal structure does not indicate any obvious role for these last three mutations, Asn^{30L} and Asp^{65H} are two of only three residues in the entire 48G7 Fab structure that have backbone (ϕ, ψ) angles that lie in disallowed regions of the Ramachandran plot. A detailed understanding of the structural and functional consequences of the somatic mutants on hapten binding and catalysis must await the high resolution structure of the germline Fab fragment, coupled with an analysis of individual somatic mutants. The germline Fab and each of the half-germline Fabs have been crystallized in the same space group with similar unit cell parameters.

It is interesting to compare the above results to other studies of somatic mutations where structural information is available. For example, analysis of the positions of the somatic mutants in a crystal structure of the monoclonal antibody (ANO2), which binds a dinitrophenyl spin label, shows that most of the mutated residues are not in direct contact with hapten (although the effect of these mutations on hapten binding was not determined) (*34*). The affinity maturation of the 2-phenyl-5-oxazolone (pheOx)–specific response by V_κOx1 and V_HOx1 with antibodies is a combination of hapten proximal and hapten distal mutations (*35*, *36*). The secondary response phenylarsonate-specific antibody 36-71 has 16 somatic mutations in the variable region genes; none directly contact the hapten (*17*). This antibody has an affinity for hapten that is ~200 times that of similar antibodies from the primary response which use the same germline genes. Comparison of the crystal structure of 36–71 to that of the primary response antibody R19.9 (which uses the same germline V genes) showed that three of six V_H and three of ten V_κ somatic mutations in 36-71 significantly perturbed the structure. Directed mutagenesis studies indicated that most of the affinity maturation in the heavy chain is caused by two mutations; neither contact the hapten (*17*). In addition, affinity maturation experiments in vitro have also resulted largely in selection for mutations distant from the active site (*37*).

Large improvements in hapten binding affinity and catalysis can result from somatic mutations distant from the antibody combining site. These mutations may have a conformational role in reorganizing active site residues to the optimal geometries for binding hapten. Another possibility is that germline antibodies can adopt a number of nearly isoenergetic conformations resulting from the flexibility of the six CDRs and amino acid side chains, and mutations fixed during affinity maturation act to "freeze" out the optimal hapten binding conformation. There is kinetic evidence for conformational isomers in a small subset of antibodies from the secondary and tertiary responses to 2-phenyl-5-oxazolone, and it has been suggested that these isomers may play a role in recognition of antigen by germline antibodies (*38*). Either way, it is likely that both sequence and configurational diversity play key roles in the affinity maturation of 48G7. Indeed, the combinatorial optimization of antigen binding by antibodies may be facilitated by a variable region structure that is highly sensitive to somatic mutations throughout the hypervariable loops.

An examination of the phylogenetic history of proteins supports, in broad outline, a scenario of exon shuffling to generate novel protein frameworks followed by natural selection acting on point mutations to improve function. The recombination and selection of antibody gene segments followed by somatic hypermutation recapitulate these processes on a dramatically reduced timescale during the immune response. Analysis of the immunological evolution of the catalytic antibody 48G7 reveals that the K_d for the transition state analogue improved by $>10^4$ during affinity maturation, with a corresponding 10^2 improvement in k_{cat}/K_m. Thus, as with enzymes, improvements in transition state binding are positively correlated with improvements in catalysis. Conformational effects arising from somatic mutation appear to play an important role in this process. Our study again emphasizes the significance of mutations distant from the active site in the natural evolution of binding and catalysis (*39*) and underscores the importance of developing methods to directly select or screen libraries of random variable region mutants to isolate catalytic antibodies with enhanced catalytic rates (*13*, *40*). Further structural and functional analysis of the affinity maturation of catalytic antibodies should increase our understanding of the evolution of catalytic function both in nature and in the laboratory.

REFERENCES AND NOTES

1. S. Tonegawa, *Nature* **302**, 575 (1983).
2. S. Kim, M. Davis, E. Sinn, P. Patten, L. Hood, *Cell* **27**, 573 (1981).
3. G.W. Siskind and B. Benacerraf, *Adv. Immunol.* **10**, 1 (1969).
4. L. A. Steiner and H. N. Eisen, *J. Exp. Med.* **126**, 1143 (1967).
5. H. N. Eisen and G. W. Siskind, *Biochemistry* **3**, 996 (1964).
6. W. Gilbert, *Nature* **271**, 501 (1978).
7. L. Pauling, *Chem. Eng. News* **24**, 1375 (1946).
8. J. B. S. Haldane, *Enzymes* (Longmans Green, London, 1930).
9. P. G. Schultz and R. A. Lerner, *Science* **269**, 835 (1995).
10. R. A. Lerner, S. J. Benkovic, P. G. Schultz, *ibid.* **252**, 659 (1991).
11. A. G. Cochran and P. G. Schultz, *ibid.* **249**, 781 (1990).
12. A. C. Braisted and P. G. Schultz, *J. Am. Chem. Soc.* **116**, 2211 (1994).
13. S. A. Lesley, P. A. Patten, P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1160 (1993).
14. J. W. Jacobs, thesis, University of California, Berkeley (1989).
15. P. Carter *et al.*, *Bio/Technology* **10**, 163 (1992); H. Ulrich *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11907 (1995).
16. Expression of the resulting chimeric Fab yields approximately 2 mg of Fab per liter when cells are grown and induced with limiting phosphate in shake flasks and approximately 30 to 300 mg of Fab per liter when cells are grown in a 2-liter fermentor with limiting phosphate. Electrospray mass spectrometry confirmed that the antibody was fully processed and disulfide-linked (predicted 48G7 mass: 47,219.93 kD, measured 48G7 mass: 47,218.23 ± 1.6 kD). Reactions for the catalytic assays were initiated by the addition of 10 μl of a freshly prepared stock solution of substrate **1** in acetonitrile to 490 μl of antibody solution at 0.2 to 1.0 mg/ml in kinetics buffer (KB consists of 20 mM tris, pH 8.2, and 50 mM NaCl) at 27°C. Hydrolysis rates were measured by absorbance at 400 nm.
17. R. K. Strong, G. A. Petsko, J. Sharon, M. N. Margolies, *Biochemistry* **30**, 3749 (1991).
18. The number N, ways to distribute r point mutations over n residues, is $n!/(n-r)!r!$. On average, any given codon can mutate to codons for six distinct amino acids. Thus, there are $(6^r)n!/(n-r)!r!$ distinct r-step point mutants that can occur over n residues. There are ~54 residues in the 48G7 hypervariable regions and ~219 residues in the combined variable domains. Thus, for $r = 9$ and $n = 54$ (CDR's) or $r = 9$ and $n = 219$ ($VJ_K + VDJ_H$), there are 5.4×10^{16} or 2.7×10^{22} distinct nine single-step somatic mutants of 48G7.
19. S. A. Kauffman, E. D. Weinberger, A. S. Perelson, *Maturation of Immune Response via Adaptive Walks on Affinity Landscapes*, A. S. Perelsons, Ed. (Addison-Wesley, Santa Fe, 1988), vol. 2, pp. 349–382.
20. It is difficult to estimate the total number of B cells that arise from a single primary B cell during the 1°, 2°, and 3° responses. A generous upper limit for the size of a clone arising during 7-day primary response if we assume an 8-hour doubling time and purely exponential growth throughout the 1° response, is 10^6 cells.
21. A. C. Malmborg, A. Michaelsson, M. Ohlin, B. Jansson, E. A. K. Borrebaeck, *Scand. J. Immunol.* **35**, 643 (1992).
22. J. Foote and C. Milstein, *Nature* **352**, 530 (1991).
23. The data set for the chimeric Fab fragment of 48G7 with hapten (*45*) was collected at the Stanford Synchroton Radiation laboratory at 4°C to a maximum resolution of 1.9 Å from crystals of typical size 0.4 mm by 0.4 mm by 0.1 mm. From precession photographs, the space group was determined to be *C2*. The unit cell lengths are $a = 140.26$ Å, $b = 44.85$ Å, $c = 85.45$ Å, and $\beta = 121.55°$. The search molecule was a chimeric 4D5 antibody (*46*) with an H3 from Hye110 (*47*) and L3 from D1.3 (*48*). The search model had ~100 side chains trimmed to Cb where the sequence is different or the B values were 2× the average. Search models for rotation calculation included the chimeric 4D5 Fab with a modified "elbow" angle incremented by 10° through the range 135° to 175°. The rotation search was performed with several integration radii with the program

AMORE. The translation search was also performed with AMORE. The original search model was subjected to one cycle of rigid-body refinement with the program X-PLOR (*49*). A $2F_o - F_c$ electron density map was calculated, the quality of the map was analyzed with the program O (*50*), and amino acid side chains were built and placed in visible density. Repeated cycles of Powell minimization, slow-cooling simulated annealing, and model building reduced the original *R* factor from 38 to 21 percent (R_{free} = 28.7 percent). All the CDR loops were deleted, simulated-annealed omit maps were computed, and the loops were rebuilt into the electron density. Topology and parameter files were constructed for the hapten, which was refined with the program X-PLOR followed by Powell minimization, simulated annealing, and model building.

24. C. Chothia and A. M. Lesk, *J. Mol. Biol.* **196**, 901 (1987).
25. C. Chothia *et al.*, *Nature* **342**, 877 (1989).
26. M. B. Lascombe *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 607 (1989).
27. D. L. Van Vranken, D. Panomitros, P. G. Schultz, *Tetrahedron Lett.* **35**, 3873 (1994).
28. G. W. Zhou, J. Guo, W. Huang, R. J. Fletterick, T. S. Scanlan, *Science* **265**, 1059 (1994).
29. E. A. Padlan, G. H. Cohen, D. R. Davies, *Ann. Inst. Pasteur Immunol.* **136**, 271 (1985).
30. B. Golinelli-Pimpaneau *et al.*, *Current Biol.* **2**, 175 (1994).
31. V. A. Roberts, J. Stewart, S. J. Benkovic, E. D. Getzoff, *J. Mol. Biol.* **235**, 1098 (1994).
32. D. J. Jeske, J. Jarvis, C. Milstein, J. D. Capra, *J. Immunol.* **133**, 1090 (1984).
33. J. L. Jorgenson, U. Esser, B. F. d. S. Groth, P. A. Reay, M. M. Davis, *Nature* **355**, 224 (1992).
34. T. P. Theriault, D. J. Leahy, M. Levitt, H. M. McConnell, G. S. Rule, *J. Mol. Biol.* **221**, 257 (1991).
35. G. M. Griffiths, C. Berek, M. Kaartinen, C. Milstein, *Nature* **312**, 271 (1984).
36. P. M. Alzari *et al.*, *EMBO J.* **9**, 3807 (1990).
37. R. E. Hawkins, S. J. Russell, G. Winter, *J. Mol. Biol.* **226**, 889 (1992).
38. J. Foote and C. Milstein, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10370 (1994).
39. L. Hedstrom, L. Szilagyi, W. J. Rutter, *Science* **255**, 1249 (1992); see also C. Chen *et al.*, *EMBO J.* **14**, 2784 (1995).
40. D. Hilvert, S. H. Carpenter, K. D. Nared, M.-T. M. Auditor, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4953 (1988); J. A. Smiley and S. J. Benkovic, *ibid.* **91**, 8319 (1994).
41. E. Kabat, *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Washington, DC, 1991); _____, T. T. Wu, M. Reid-Miller, H. M Perry, K. S. Gottesman, *Sequences of Proteins of Immunological Interest* (Government Printing Office, Bethesda, MD, 1991).
42. The unmutated Balb/c germline V_κ and V_H genes used by the 48G7 hybridoma (Balb/c) were cloned as follows. Inspection of the V regions of 48G7 revealed that this hybridoma uses members of the V_κ subgroup 5 and V_H subgroup 2(C) families (*41*). The rearranged 48G7 V_κ and V_H genes, as well as the germline Balb/c MOPC41 V_κ and H10 V_H families were amplified by PCR with primers based on the 5′ and 3′ flanking sequences derived from genomic clones of MOPC41 V_κ and H10 V_H, as well as 48G7 CDR3 junctional sequences (*51*). The 5′ flanking sequences encompassing the 5′ untranslated (UT), leader, and leader-V region intron sequences were determined for subclones of the rearranged 48G7 V_κ [280 base pair (bp)] and 48G7 V_H (261 bp) genes. Sequencing of this same 5′ flanking region in subclones of the Balb/c germline MOPC41 V_κ family allowed us to unambiguously identify the germline V_κ gene used by 48G7 according to the following argument. Nine distinct sequences were identified from the MOPC41 V_κ subfamily. One germline gene was identical to 48G7 at 279 of 280 bases in the 5′ flanking region, whereas the other five germline genes differed by 7 to 17 base changes and 2 to 9 deletions. There are four differences (three missense, one silent) between the coding regions of 48G7 V_κ and this germline V_κ (Fig. 1), and we conclude that these represent somatic mutations in the light chain. The 260 bp of 48G7 V_H 5′ flanking sequences that were analyzed, encompassing 5′ UT, leader, and leader-V region intron sequences, are identical to those of the Balb/c germline gene H10, one of only three closely related members of the H10 V_H subfamily (*52*). The two other reported members of this subfamily, H4a.3 and H2b.3, contain, respectively, 23 or 28 base changes and three or two deletions relative to 48G7 in the 5′ flanking region (*52*). We therefore conclude that H10 is the Balb/c germline gene that was utilized by 48G7 and that the six replacements and one silent change in the 48G7 V_H coding region are somatic mutations.
43. Hapten **3** was coupled to the amino terminus of resin-bound peptide **4** (ASQYFPSQAHGSK) with the use of HBTU and HOBT. After acetylation, cleavage from the resin, and deprotection, the products (**5, 6**) were purified on a reversed phase HPLC (Vydak) and verified by mass spectrometry. The peptide-hapten conjugate was coupled to the surface by standard procedures until ~123 RU's were bound. Negative control surfaces were generated by coupling **6** under identical conditions. Surfaces were inactivated by treatment with ethanolamine. Recombinant Fab was affinity purified (*15*), concentrated to 3 to 10 mg/ml, and further purified by FPLC on a Superdex 200 column to remove aggregates. Rigorous size purification was done to remove a small amount of a high avidity component in the Fab preparations which presumably corresponds to multimeric Fab aggregates (not detectable on silver-stained protein gels). The G^L-G^H Fab FPLC peak was further concentrated with Centricon-30 membranes to ~25 mg/ml for Scatchard analysis, in which injections over flow cells coupled with (i) peptide-hapten **5** or (ii) acetylated peptide **6** were done under conditions where binding reached equilibrium. Specific binding at equilibrium was calculated by subtracting equilibrium binding in (ii) from that in (i). Twofold dilutions of Fab covering the following ranges were used: $48G7^L$-$48G7^H$ 10 to 160 μM; $48G7^L$-G^H 30.7 to 0.03 μM; G^L-$48G7^H$ 33.7 to 0.33 μM; G^L-G^H 405.6 to 0.396 μM. For kinetic studies, each Fab was diluted in KB to the working concentrations indicated in Fig. 4. These solutions were then injected (20 μl of Fab at 5 μl/min) over the surface coupled with (i) TSA **5** or (ii) the negative control **6**. In each case the presence of 1 mM **3** completely inhibited specific binding. The k_{off}'s for $48G7^L$-$48G7^H$, G^L-$48G7^H$, and $48G7^L$-G^H were determined in separate runs by filling loop 1 with a high concentration of Fab (> 100 μg/ml) and loop 2 with KB + 1 mM TSA **3**. The contents of loop 1 were injected at 30 μl/min with the contents of loop 2 being injected at 30 μl/min, precisely as the dissociation phase began in order to eliminate rebinding during the dissociation phase. The k_{on} and k_{off}'s for $48G7^L$-$48G7^H$ were then calculated by nonlinear curve fitting (BIA evaluation II software). The k_{on}'s for G^L-$48G7^H$ and $48G7^L$-G^H were too great to be resolved on the instrument, but could be calculated as k_{off}/K_D.
44. The association and dissociation kinetics were determined and could be fit, for 48G7, to the simple model in which Fab and **3** reversibly associate to form the Fab·**3** complex. The rapid approach to equilibrium by the germline and half-germline antibodies is a consequence of the high Fab concentrations required to detect specific binding. Although the association rate constants for the half-germline antibodies could not be kinetically resolved on the instrument, k_{on} values could be calculated from the measurable apparent k_{off} and equilibrium K_d. The low affinity of the germline antibody precludes precise kinetic measurement of either k_{on} or k_{off}, and only allows us to place bounds on the rate constants. The validity of the Scatchard analysis is supported by the agreement between the K_d derived by Scatchard and kinetic analyses.
45. Purified $48G7^L$-$48G7^H$ (48G7) Fab fragment was concentrated to 15 mg/ml in a buffer consisting of 100 mM NaCl, 10 mM tris, pH 8.0, 1.0 mM methionine, 0.5 mM EDTA, 0.3 mM sodium azide; hapten, when present, was at a concentration of 1 mM. Crystallization screens for the Fab fragment (*53*) were set up in hanging drops at 4°C, with each drop consisting of 2 μl of protein or protein and hapten mixed with 2 μl of mother liquor, and each well containing 1 ml of mother liquor. Crystals grew within a few days, and continued to appear for several weeks. Several different crystal morphologies were seen, but invariably the only crystal type that diffracted adequately for data collection were platelike crystals having ill-defined faces and long-axis dimensions in excess of 0.5 mm. The following optimal crystallization conditions were determined for 48G7 Fab (15 mg/ml) with hapten (1 mM): 100 mM sodium acetate, 100 mM sodium cacodylate, pH 7.0, 27 percent PEG 3300.
46. C. Eigenbrot, M. Randal, L. Presta, P. Carter, A. A. Kossiakoff, *J. Mol. Biol.* **229**, 969 (1993).
47. E. A. Padlan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5938 (1989).
48. T. N. Bhat, G. A. Bentley, T. O. Fischmann, G. Boulot, R. J. Poljak, *Nature* **34**, 483, (1990).
49. A. T. Brünger, X-PLOR (New Haven, CT, 1992).
50. T. A. Jones and K. Kjeldgaard, "O" Computer Graphics Program.
51. The following oligonucleotides, used for polymerase chain reaction (PCR) amplification and analysis of the described constructs, were synthesized and purified by standard methods.
    1) 5′-AAGGAACTCGAGCGTGACCAATCCTAACTGCTTCTTAATAAT-3′
    2) 5′-CGTAGAATTCTCAGAGCAGCCCAGCCTCTCATTTC-3′
    3) 5′-ATCGAATTCGAATGTCCTTATGTAAGAAA-3′
    4) 5′-CGTACTCGAGGACTGTGAGAGTGGTGCCTTGGCC-3′
    5) 5′-GACTGTGAGAGTGGTGCCTTGGCCCCAGTAGATACCATAGTAGCTAGCACAGTAATA-3′
    6) 5′-GACCGATGGGCCCTTGGTGGAGGCTGAGGAGACTGTGAGAGTGGTGCC-3′
    7) 5′-CTTGTGAAGCCAGGG-3′
    8) 5′-AAAGACACCTATATG-3′
    9) 5′-ATTGGAAGGATTGATCCTGCGAATGTT-3′
    10) 5′-GCTACAAACGCGTACGCTGATATCCAGATGACCCAGTCTCCATCCTCC-3′
    11) 5′-CAGTTTGGTACCTCCACCGAACGTCCGAGGATAACTAGCATATTGCAGACA-3′
    12) 5′-AAAGGGGAATTCTCCACCGAACGTCCGAGGATA-3′
    13) 5′-GCTACAAACGCGTACGCTCAGGTTCAGCTGCAGCAGTCTGGGGCAGAGCTTGTG-3′

    The rearranged 48G7 VJ_κ and VDJ_H genes were PCR amplified from 48G7 genomic DNA with oligos 1+12 and 3+4, respectively. The germline BALB/c MOPC41 V_κ and subgroup 2(C) V_H subfamilies were PCR-amplified from BALB/c tail DNA with oligos 1+2 and 5+13, respectively. The PCR fragments were subcloned into pBS-KS for sequencing and further manipulations. Subclones (300) of subgroup 2(C) inserts were screened for H10 homologous sequences with oligos 7, 8, and 9. The 48G7 VJ_κ and VDJ_H were reconfigured for cloning into pMY61 by PCR amplification with oligos 10+11 and 6+13 respectively. The germline V_κ and V_H genes were recombined with the 48G7 J_κ and DJ_H regions and reconfigured for cloning into pMY61 by PCR amplification with 10+11 and 5+6+13, respectively. The VJ_κ and VDJ_H inserts were digested with Mlu I and Kpn I and Mlu I and Apa I, respectively, and cloned into pDEI440, a derivative of pMY61 containing an M13 origin of replication.
52. C. Schiff, M. Milili, I. Hue, S. Rudikoff, M. Fougereau, *J. Exp. Med.* **163**, 573 (1986).
53. J. Jancarik and S.-H. Kim, *Appl. Crystallogr.* **24**, 409 (1991).

54. We thank M. Davis for use of the BIAcore; D. Henner, S. Bass, and B. Snedecore (Genentech) for the plasmids pMY-55, -60, and -61; D. King for performing electrospray mass spectrophotometry; Stanford Synchrotron Radiation Laboratory for data collection time and H. Morimoto (National Tritium Labeling Facility) for assistance synthesizing tritiated hapten **3**. Supported in part by the National Institutes of Health (RO1-AI24695), the Howard Hughes Medical Institute, a Damon Runyon–Walter Winchell Cancer Research Fund Fellowship (P.A.P.), a Howard Hughes Predoctoral Fellowship (P.L.Y.), and the Irvington Institute for Medical Research (J.J.B.). The coordinates have been deposited in the Brookhaven Protein Data Bank (PDB, 11GAF).

1 September 1995; accepted 26 December 1995