

the membrane is of infinite extent. The Euler angles and protein-membrane distance have errors of less than  $10^\circ$  and  $2 \text{ \AA}$ , respectively, as given by the standard error of propagation for nonlinear least squares algorithms. The data for I78C-sl were not included in the fit because they are clearly anomalous (Table 1). Because this mutant has only 7% of the enzymatic activity of the wild type, EPR data are not useful. All fitting was done with MATLAB (Mathworks, Cambridge, MA).

18. Because the Crox-nitroxide spin exchange happens as fast as Crox diffuses to the nitroxide (9), the possibility that the presence of the membrane slows the rate of spin relaxation by slowing the rate of diffusion-limited Crox-nitroxide encounters was also considered, but the calculation (available from the authors on request) shows that diffusional effects only occur if the membrane is  $<3 \text{ \AA}$  from the spin label, and thus diffusional effects are not responsible for the deviation of  $\Phi$  from unity measured when the spin label is tens of angstroms away from the membrane.

19. The electrostatic potentials for bvPLA<sub>2</sub> bound to DTPM vesicles ( $\psi_{E,M}$ ) as indicated in Fig. 1 and for the protein ( $\psi_E$ ) and membrane ( $\psi_M$ ) were calculated as previously described [N. Ben-Tal, B. Honig, R. M. Peitzsch, G. Denisov, S. McLaughlin, *Biophys. J.* **71**, 561 (1996)]. Each leaflet of the membrane bilayer consisted of 360 hexagonally packed DTPM lipids, and it was assumed that  $\text{Ca}^{2+}$  and one DTPM were bound in the enzyme's active site. For each spin label, the potential values in the aqueous phase within  $5 \text{ \AA}$  of the nitroxide nitrogen were used to calculate the average potential difference ( $\psi_{E,M} - \psi_E - \psi_M$ ). The calculated potential differences (residue and average  $\pm$  SD in millivolts) are: 2,  $-26 \pm 15$ ; 13,  $-7 \pm 5$ ; 14,  $-16 \pm 10$ ; 15,  $-6 \pm 8$ ; 23,  $-12 \pm 6$ ; 24,  $-9 \pm 3$ ; 51,  $-2 \pm 3$ ; 53,  $0 \pm 1$ ; 66,  $3 \pm 1$ ; 78,  $-9 \pm 3$ ; 85,  $-6 \pm 2$ ; and 92,  $-3 \pm 1$ .

20. B. Honig and A. Nicholls, *Science* **268**, 1144 (1995).  
 21. T. Thuren *et al.*, *Biochemistry* **23**, 5129 (1984).  
 22. M. K. Jain, B.-Z. Yu, J. Rogers, G. N. Ranadive, O. Berg, *Biochemistry* **30**, 7306 (1991); B.-Z. Yu *et al.*, *ibid.* **36**, 3870 (1997).  
 23. F. Ghomashchi *et al.*, *ibid.*, in press.  
 24. N. Din *et al.*, *Mol. Microbiol.* **11**, 747 (1994).  
 25. P. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991); E. A. Merritt and D. J. Bacon, *Methods Enzymol.* **277**, 505 (1997).  
 26. This work was supported by NIH grants HL36235 (M.H.G.), GM32681 (B.H.R.), and The Center for Environmental Health P30 ESO7033 (B.H.R.). D.M. is a Helen Hay Whitney Postdoctoral Fellow. We are grateful to J. M. Schurr, S. McLaughlin, and L. J. Slutsky for discussion of theory; to T. Lybrand and E. Adman (NIH University of Washington Center grant P30ES07033) for help with graphics; and to A. Beth for  $^{15}\text{N}$  spin labels.

## Immunological Origins of Binding and Catalysis in a Diels-Alder Antibody

Floyd E. Romesberg,\* Ben Spiller,\* Peter G. Schultz,†  
 Raymond C. Stevens†

The three-dimensional structure of an antibody (39-A11) that catalyzes a Diels-Alder reaction has been determined. The structure suggests that the antibody catalyzes this pericyclic reaction through a combination of packing and hydrogen-bonding interactions that control the relative geometries of the bound substrates and electronic distribution in the dienophile. A single somatic mutation, serine-91 of the light chain to valine, is largely responsible for the increase in affinity and catalytic activity of the affinity-matured antibody. Structural and functional studies of the germ-line precursor suggest that 39-A11 and related antibodies derive from a family of germ-line genes that have been selected throughout evolution for the ability of the encoded proteins to form a polyspecific combining site. Germ line-encoded antibodies of this type, which can rapidly evolve into high-affinity receptors for a broad range of structures, may help to expand the binding potential associated with the structural diversity of the primary antibody repertoire.

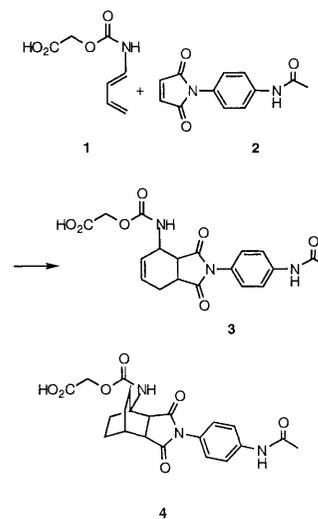
The immune system solves the problem of molecular recognition by generating a large library of structurally distinct antibodies and amplifying those with the requisite binding affinity and specificity in an affinity-based selection. By programming this system with chemical information about a reaction mechanism—for example, the structure of a putative transition state—one can examine the evolution of both binding energy and catalytic function (1). Functional and structural analysis of this process can provide insights into both the molecular basis for the remarkable efficacy of this combinatorial system and the mechanisms

by which binding energy can be used to lower the activation energies of reactions (1–5). We now describe one such study of the antibody 39-A11 (6), which catalyzes a Diels-Alder reaction, a widely used and mechanistically well studied reaction in organic chemistry, but one that is rarely found in biological systems. The three-dimensional x-ray crystal structures of the 39-A11 Fab-hapten complex and of the germ-line precursor have been determined, and the immunological origins of this and related antibodies have been characterized.

Antibody 39-A11 was generated to the bicyclo[2.2.2]octene hapten 4, a mimic of the boatlike transition state of the Diels-Alder reaction. This antibody catalyzes the cycloaddition reaction of diene 1 and dienophile 2 to give the Diels-Alder adduct 3 (Scheme 1) (6). Structurally related haptens have been used to generate other antibodies that catalyze Diels-Alder reactions, suggesting that this is a relatively general design strategy (7, 8). Antibody 39-A11 was cloned and expressed as a humanized chi-

meric Fab (9), and the structure of the complex of the recombinant 39-A11 Fab fragment and hapten 4 was determined at  $2.4 \text{ \AA}$  resolution (Fig. 1 and Table 1).

Well-defined density for the hapten was observed in the  $1F_o - 1F_c$  omit map (Fig. 1). The hapten is bound in a cleft  $\sim 9 \text{ \AA}$  wide and  $\sim 12 \text{ \AA}$  deep, with  $\sim 194 \text{ \AA}^2$  of the hapten surface (79% of the total solvent-accessible surface excluding the linker arm) buried within the Fab. There are 89 van der Waals interactions and two hydrogen bonds between the hapten and antibody, with most of these contacting the heavy chain. The bicyclo[2.2.2]octene moiety of hapten 4, which corresponds to the cyclic  $4+2 \pi$  electron system of the transition state, is buried in a hydrophobic pocket, free of solvation. The walls of this cavity consist of the side chains of residues Phe<sup>H100b</sup> [antibody nomenclature described in (20)], Asn<sup>H35</sup>, Trp<sup>H47</sup>, Val<sup>L91</sup>, Pro<sup>L96</sup>, Gly<sup>H33</sup>, Trp<sup>H50</sup>, Ala<sup>H95</sup>, and Arg<sup>H100</sup> (where H and L represent heavy and light chains of the antibody, respectively). The carbonyl oxygen of the carbamate moiety at the bridgehead position of 4 (the C1 substituent in



Scheme 1

F. E. Romesberg and P. G. Schultz, Howard Hughes Medical Institute and the Department of Chemistry, University of California, Berkeley, CA 94720, USA, and Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

B. Spiller and R. C. Stevens, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, and the Department of Chemistry, University of California, Berkeley, CA 94720, USA.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed.

the diene) is hydrogen-bonded through a water molecule to Nε1 of Trp<sup>H50</sup>. The succinimido moiety of **4**, which corresponds to the maleimide group of the dienophile, is involved in a  $\pi$ -stacking interaction with Trp<sup>H50</sup>. The *N*-phenyl substituent is packed against the backbone of residues Gly<sup>H33</sup> and Glu<sup>H96</sup> and the methylene groups of Arg<sup>H97</sup>. The succinimido carbonyl group anti to the carbamate moiety of **4** is hydrogen-bonded to the side chain amide of Asn<sup>H35</sup>, which is oriented by a hydrogen bond between the carboxamide oxygen and the side chain of Trp<sup>H47</sup>.

The x-ray crystal structure of the Fab-hapten **4** complex suggests that antibody 39-A11 binds the diene and dienophile in a reactive orientation and reduces translational and rotational degrees of freedom (10, 11). The dienophile is oriented by hydrogen-bonding and  $\pi$ -stacking interactions with the maleimide ring. The diene is bound in a

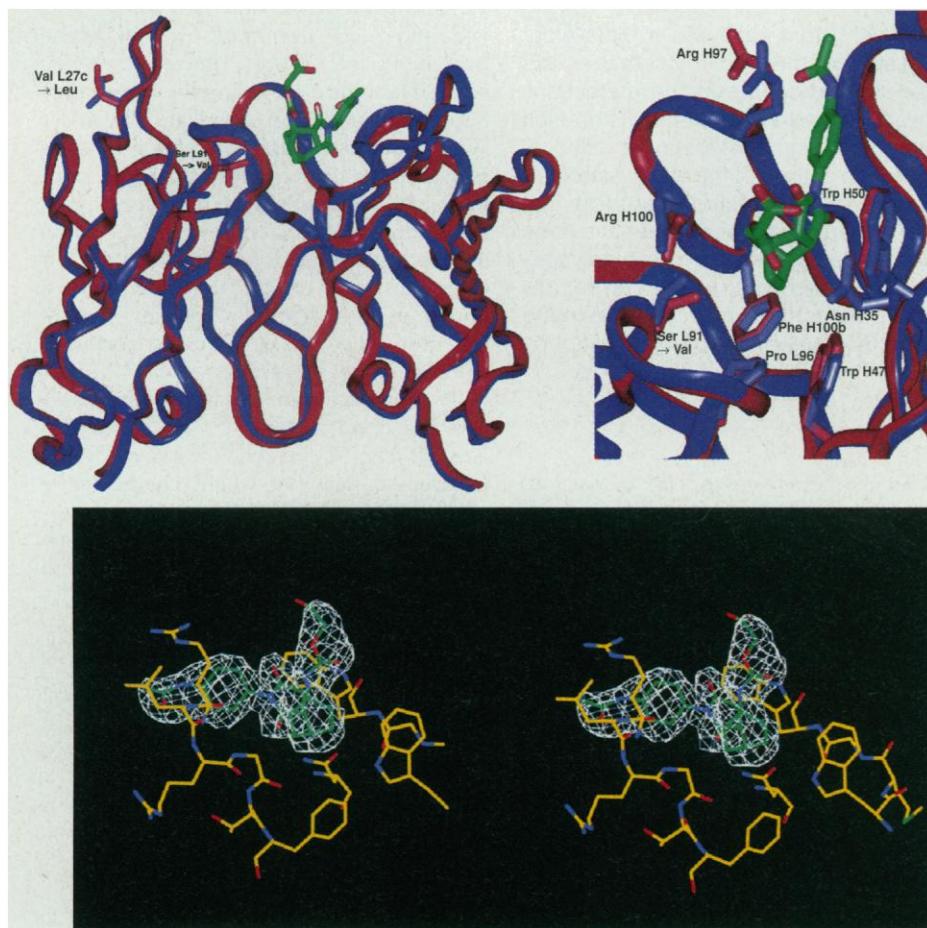
hydrophobic pocket in close proximity to the dienophile, with the position of the carbamate substituent fixed by a water-mediated hydrogen bond to Trp<sup>H50</sup> (Fig. 1). Although the carbon-carbon bond lengths of bicyclo[2.2.2]octene are shorter than those that would be present in either a synchronous or nonsynchronous transition state, it appears that both transition states can be accommodated in the active site. In addition to proximity effects, the energetics of the antibody-catalyzed reaction may be influenced by a hydrogen bond between the side chain carboxamide group of Asn<sup>H35</sup> and the maleimide group. This interaction should render the olefin more electron deficient and, as a result, a more reactive dienophile (Fig. 1) (12). In the instance of an asymmetric dienophile, containing only one electron-withdrawing carbonyl group, this interaction might be expected to enhance the formation of the disfavored regioisomer of the Diels-

Alder adduct.

The structure of the 39-A11-hapten **4** complex provides an explanation for the  $\sim 1000$  times greater binding affinity of 39-A11 for hapten **4** than for product **3**. Conformational constraints imposed by the [2.2.2]-bicyclic framework lock the cyclohexene ring of the hapten into a boatlike geometry distinct from that of the product. These constraints result in less favorable van der Waals and hydrogen-bonding interactions between active site residues and substituents on product **3** than between active site residues and hapten **4**, for example, the water-mediated hydrogen bond between Trp<sup>H50</sup> and the carbamate moiety of the linker. The crystal structure also suggests that mutation of the active site residues Pro<sup>L96</sup> and Val<sup>L91</sup> in 39-A11 to residues with increased hydrophobic surface area might lead to an increase in rate as a result of improved packing interactions with the kinetically favored *endo* transition state (the structure of hapten **4** accommodates both *endo* and *exo* transition states). Indeed, three out of six site-directed mutations of these residues (Val<sup>L91</sup> to Tyr, and Pro<sup>L96</sup> to Phe or Tyr) were associated with about 5- to 10-fold increases in the catalytic rate constant ( $k_{cat}$ ).

The structure of antibody 39-A11 can be compared with those of other proteins that catalyze pericyclic reactions. In the case of the antibody 1F7 (5) and the *Bacillus subtilis* (13), *Escherichia coli* (14), and yeast (15) chorismate mutases, which catalyze the Claisen rearrangement of chorismate to prephenate, the protein also appears to organize the substrate into a reactive conformation by a network of hydrogen-bonding and van der Waals interactions (10, 11). In the enzyme active sites, there are also hydrogen bonds to the enol ether oxygen of chorismate, which have been postulated to stabilize the developing charge at this center in the transition state (16). Nuclear magnetic resonance and x-ray crystallographic analyses of monoclonal antibody AZ-28, which catalyzes a related pericyclic rearrangement (oxy-Cope), indicate that this antibody also stabilizes the conformationally restricted cyclic transition state (3). In addition, AZ-28 may increase the reaction rate by enhancing the extent of electron density on the hydroxyl substituent of the substrate through hydrogen-bonding interactions. Thus, it appears that catalysis of these pericyclic transformations involves both restriction of rotational and translation entropy in the substrate as well as hydrogen-bonding interactions that modulate electron densities on key substituents in the transition state.

Analysis of the germ-line precursor to antibody 39-A11 provides an opportunity



**Fig. 1.** (Left) Ribbon superposition of the variable regions of the germ-line Fab without hapten (purple) and the mature Fab-hapten **4** complex (red). The side chains of the somatic mutation sites are indicated for the germ-line and mature antibodies: Val<sup>L27c</sup>→Leu and Ser<sup>L91</sup>→Val. (Right) Close-up view of the Diels-Alder 39-A11 active site with the transition state analog bound to the mature form of the antibody. The mature antibody is shown in red, and the germ-line antibody without hapten in purple. No significant differences are apparent between the two structures. The hapten molecule is shown with carbon atoms in green, oxygen in red, and nitrogen atoms in blue. (Bottom) Stereoview of the  $1F_o - 1F_c$  electron density surrounding the transition state analog ( $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively).

to examine the evolution of this biological catalyst in a combinatorial system that reflects features of the natural evolutionary process. The  $V_L$  gene for 39-A11 exhibits three nucleotide differences (one of which is silent) relative to its germ-line precursor  $V_{\kappa}$ 1A (Fig. 2) (17); the V region is joined in frame to  $J_{\kappa}$ 1, resulting in a proline at junctional position 96. The  $V_H$  gene differs by one nucleotide (which is silent) from its germ-line precursor VMS9, a member of the VGAM3.8  $V_H$  family (18). There may also be one difference in the last codon of the gene, but this nucleotide is likely not encoded by the variable region and therefore probably does not represent a somatic mutation; D region DSP2.2 (19) and  $J_H$ 4 (20) genes were used in unmutated form (Fig. 2). This analysis indicates that affinity maturation of antibody 39-A11 results in only two somatic mutations: a Val→Leu substitution at position 27c in CDRL1, and a Ser→Val substitution at position 91 in CDRL3. The functional consequences of affinity maturation on binding affinity and catalysis were determined by expressing and characterizing the germ-line antibody as well as the individual somatic mutants. The dissociation constant ( $K_d$ ) of the germ-line antibody for hapten 4 is  $379 \pm 16$  nM (21). The two somatic mutations result in a 40-fold increase in binding affinity ( $K_d = 10 \pm 0.3$  nM) of the mature antibody 39-A11 for hapten 4, with virtually all of this increase associated with the somatic mutation at position L91. The effects of the somatic mutations on  $k_{cat}$  and the Michaelis constant ( $K_m$ ) parallel their effects on binding affinity. The  $k_{cat}$ ,  $K_m$ (1), and  $K_m$ (2) values for the recombinant Fab 39-A11 are  $0.67$  s<sup>-1</sup>,  $1200$  μM, and  $740$  μM, respectively, whereas the corresponding values for the germ-line antibody are  $0.17$  s<sup>-1</sup>,  $1400$  μM, and  $450$  μM (21). The effects of affinity maturation on catalysis, which are reflected primarily in the values of  $k_{cat}$  and  $K_m$ (1), again result largely from the Ser<sup>L91</sup>→Val somatic mutation.

To probe the structural consequences of affinity maturation, we determined the three-dimensional structure of the germ-line antibody Fab at 2.1 Å resolution (Fig. 1 and Table 1). The structures of the unliganded germ-line antibody and the 39-A11-hapten complex are very similar. The root-mean-square deviations for the Cα positions of  $V_H$  and  $V_L$  are 0.45 and 0.51, respectively; the overall deviation for the variable domain Cα positions is 0.71. The only relatively large structural change is in the position of the poorly defined side chain of Phe<sup>L87</sup>, which is distant from the active site. Smaller differences are apparent at the site of the

Ser<sup>L91</sup>→Val somatic mutation (the distance from Val<sup>L91</sup> to the unsubstituted bridgehead carbon of 4 is 4 Å) and the nearby residue His<sup>L34</sup>. Nonetheless, a comparison of the two structures indicates that neither somatic mutation nor ligand binding results in substantial structural or conformational changes in the active site.

Thus, the affinity maturation of antibody 39-A11 presents an opportunity to examine a solution to the problem of molecular recognition markedly different from that involving the esterolytic antibody 48G7,

which we previously described (2). In the latter instance, nine somatic mutations, none directly contacting the hapten, contribute additively to the 30,000-fold increase in affinity for the nitrophenyl phosphonate transition state analog ( $K_d = 4$  nM). Structural changes occur on binding of hapten to the germ-line antibody that result in enhanced antibody-hapten complementarity. These structural changes were further optimized by affinity maturation, resulting in “lock and key binding” of hapten to the mature antibody. In contrast, the germ-line

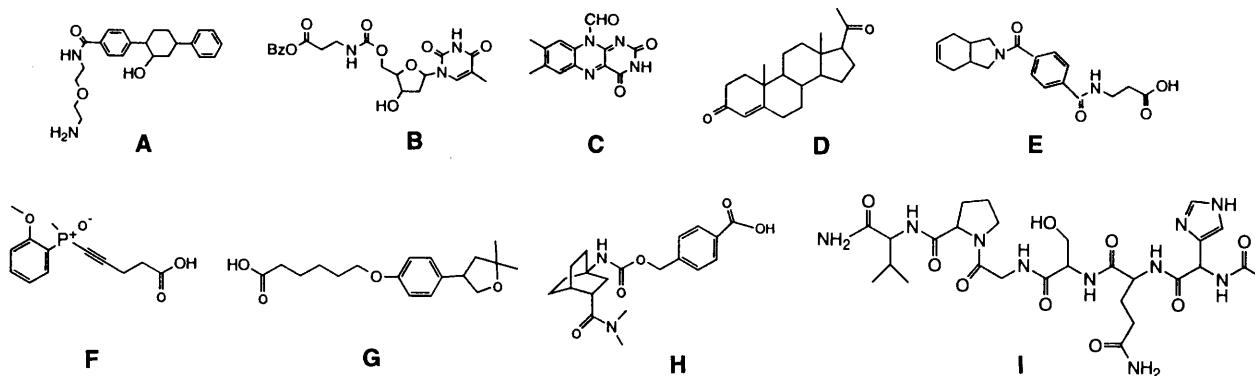
**Table 1.** Data collection and refinement statistics for the mature (with hapten) and germ-line (without hapten) Fabs (31).

Item	Mature	Germ line
Space group	$P2_12_12_1$	C2
Unit cell dimensions (Å)	$a = 69.3$ $b = 96.8$ $c = 172.7$	$a = 158.9$ $b = 49.3$ $c = 145.0$ $\beta = 108.7^\circ$
Refinement resolution (Å)	20.0–2.4	20.0–2.1
Observations ( $n$ )	243,063	423,134
Unique reflections ( $n$ )	44,113	62,308
$R_{sym}$ ( $I$ ) <sup>*</sup> (%)	4.7	4.5
$R_{cryst}$ <sup>†</sup> (%)	21.1	22.9
$R_{free}$ <sup>‡</sup> (%)	29.1	29.2
Root-mean-square deviation		
Bond lengths (Å)	0.007	0.008
Bond angles (degrees)	1.351	1.383
Completeness to refined resolution	99.5 (99.8) <sup>§</sup>	98.5 (95.4) <sup>§</sup>

<sup>\*</sup> $R_{sym} = 100 \times \sum_n \sum_i |I_{n,i} - \bar{I}_n| / \sum_n \sum_i I_{n,i}$  for the intensity  $I$  of  $i$  observations of reflection  $h, l, j$ ,  $\bar{I}_n$  is the mean intensity of the reflection. <sup>†</sup> $R_{cryst} = 100 \times \sum |F_{obs} - F_{calc}| / \sum |F_{obs}|$ , where  $F_{obs}$  and  $F_{calc}$  are the observed and calculated structure factors, respectively. <sup>‡</sup> $R_{free}$  is the same as  $R$  factor, but is calculated from the 10% of the reflection data excluded from refinement. <sup>§</sup>Highest resolution bin given in parentheses.

	<u>CDRL1</u>		
$V_{\kappa}$ 1A	DVVMITQTPLSLFLVSLGDAQSISCS	RSSQSLVHNSGNTYLH WYLQKPGQSPKLLIY	
39-A11	-----L-----	-----L-----	
DB3	-----I-----N-----	-----I-----	
TE33	--L-----	K----IV--S----FE-----	
	<u>CDRL2</u>	<u>CDRL3</u>	
$V_{\kappa}$ 1A	KVSNRFS	GVPDRFSGSGSGIDFTLKISRVEAEDLGVYFC	SQSTHVP
39-A11	-----	-----V-----	P T-FGGGTLKLEIK
DB3	-----Y-----	-----I-----	-----
TE33	-----	-----Y- F-GS-I- F - -S-----	-----
	<u>CDRH1</u>	<u>CDRH2</u>	
VMS9	QIQLVQSGPELKKPGETIVKISCKASGYTFT	NYGMN WVKQAPGKGLKWMG	WINIYTGEPYADDFKG
39-A11	-----	-----	-----
DB3	-----A-----	-----V-----E-----E-----	-----I-----V-----
TE33	-----T-----R-----	T--S--T--F--	-----S-V-----
	<u>CDRH3</u>		
VMS9	RFAFSLETSASTAYLQINLNKEDTATYFC		
39-A11	-----	-----VQ AERLRRITFDY WGAGTIVIVSS	
DB3	-----E-----	-----TR GDYVMYFDV-----	
TE33	-----W-----	-----ARR . . . . WYFDV --T-----	

**Fig. 2.** Sequences of the  $V_H$  and  $V_L$  regions of the structurally related antibodies 39-A11, DB3, and TE33, and of the germ-line precursors VMS9 and  $V_{\kappa}$ 1A, respectively.  $V_{\kappa}$ 1A and VMS9 were identified as likely  $V_L$  and  $V_H$  germ-line candidates, respectively, through a homology search of the Kabat database (20). On the basis of their published sequences, 5' PCR primers were designed to anneal to upstream untranslated DNA of the  $V_{\kappa}$ 1A and VMS9 genes. PCR amplification with 39-A11 hybridoma DNA as template in conjunction with 3' light and heavy J region-specific primers yielded several clones of both chains, which were then sequenced. In each instance, the flanking regions were shown to be identical to the published  $V_{\kappa}$ 1A and VMS9 sequences over a region of hybridoma DNA sufficiently large to identify the rearranged gene (~500 and 400 nucleotides for the light and heavy chain genes, respectively). 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. Dashes represent residues identical to the corresponding  $V_{\kappa}$ 1A or VMS9 sequence.



**Fig. 3.** Structures of ligands used in binding assays with 39-A11 and its germ-line precursor.

precursor to antibody 39-A11 appears to be a better start point—only one somatic mutation in the combining site is required to bind hapten with an affinity similar to that of 48G7. This difference may be a consequence of the hydrophobic nature of hapten 4; alternatively, antibody 39-A11 may have evolved from a relatively polyspecific combining site that was selected in the germ-line repertoire for its ability to bind a structurally diverse array of antigens. Therefore, we assayed binding of both 39-A11 and its germ-line precursor to a panel of chemically defined, structurally diverse haptens conjugated to bovine serum albumin (BSA) (Table 2 and Fig. 3).

Both antibodies bind nine haptens containing a broad range of hydrogen-bonding, charged, and hydrophobic groups [in contrast, the nitrophenyl phosphonate-specific antibody 48G7 showed no cross-reactivity when screened against the same panel of hapten-BSA conjugates (2)]. Comparison of these affinities with those for hapten 4 provides a measure of the polyspecificity. The germ-line precursor of 39-A11 shows affinities for haptens A through I that are roughly within an

order of magnitude to that for its own hapten 4, whereas the affinity of 39-A11 for haptens A through I. These results suggest that the germ line-encoded antibody is polyspecific and can be selected for clonal expansion and subsequent affinity maturation by a wide variety of antigens, including those shown in Fig. 3. This polyspecificity may be general to several germ line-encoded antibodies and may have been selected for by the immune system to provide a mechanism for rapid generation of antibodies of moderate to high affinity for a broad range of antigens.

Three other antibodies (22–24) that were clonally selected on the basis of their intrinsic affinities for markedly different ligands use  $V_H$  and  $V_L$  chains highly homologous to those of 39-A11. Antibodies DB3 (22), TE33 (23), and IE9 (24) were raised against progesterone, a 16-amino acid peptide, and a hexachloronorborene derivative, respectively. The three-dimensional crystal structures of DB3 (22) and TE33

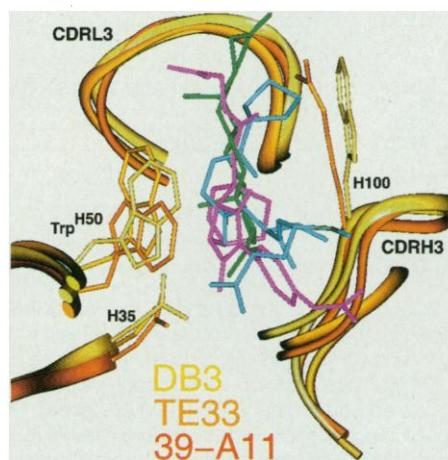
(23) have been solved and, together with the structure of 39-A11, make possible a detailed structural analysis of how an antibody of limited diversity is able to bind a variety of structurally distinct antigens (Fig. 4). Antibody DB3 is specific for progesterone and most likely comprises  $V_{\kappa}1A$ ,  $J_{\kappa}1$ , VMS9, and  $J_H4$  (25); it also shows cross-reactivity with various structurally related progesterone analogs (26). Both DB3 and IE9 show some cross-reactivity (24). Antibody TE33 is specific for the cholera toxin peptide VEVPGSQHIDSQKKA, and most likely comprises  $V_{\kappa}1C$ ,  $J_{\kappa}4$ , V264 (a member of the VGAM3.8 family), and  $J_H1$  (27). Most of the differences in sequence among these antibodies are located in CDRH3 and are not germ line-encoded (Fig. 2). All three antibodies use a light chain variable region encoded by the  $V_{\kappa}1$  gene, which is common to a relatively large population of antibodies that bind a large number of antigens including proteins, DNA, steroids, peptides, and small haptens (17).

The combining sites of 39-A11, DB3, and TE33 present a large, highly conserved binding surface formed predominantly by the CDRL1, CDRL3, and CDRH3 hypervariable loops; the CDRH3 and CDRL3 loops together with  $\text{Trp}^{H50}$  form a deep hydrophobic binding pocket (Fig. 4). A more shallow region of the binding pocket is dominated by  $V_L$  contacts, with CDRL3 providing the floor and CDRL1 bordering the pocket. In antibodies DB3 and 39-A11,  $\text{Trp}^{H50}$  and residue H100 in the CDRH3 loop sandwich the hapten, providing critical hydrogen-bonding or hydrophobic contacts that define opposite walls of the deep binding pocket— $\text{Trp}^{H100}$  in DB3 packs with the central nonpolar region of the steroid, and  $\text{Arg}^{H100}$  and  $\text{Trp}^{H50}$  in 39-A11 provide key hydrophobic and hydrogen-bonding interactions with hapten 4. In these same two antibodies, the amide side chain of  $\text{Asn}^{H35}$  at the bottom of the binding pocket is positioned to donate a key hydrogen bond to the ligand. The con-

**Table 2.** Dissociation constants for the binding of ligands to antibody 39-A11 and its germ-line precursor.

Ligand	$K_d$ ( $\mu\text{M}$ )	
	Germ line	39-A11
4*	$0.379 \pm 0.016$	$0.010 \pm 0.003$
A*	$3.10 \pm 0.01$	$5.02 \pm 0.07$
B*	$2.4 \pm 0.01$	$4.9 \pm 0.2$
C*	$7.9 \pm 0.6$	$8.4 \pm 1.2$
D*	$0.40 \pm 0.01$	$0.40 \pm 0.01$
E†	5	5
F†	20	20
G†	10	10
H†	10	10
I*	$0.10 \pm 0.01$	$0.11 \pm 0.02$

\*Determined by fluorescence quenching (21). †Determined by enzyme-linked immunosorbent assay (32).



**Fig. 4.** Superposition of the CDRL3 and CDRH3 loops of antibodies DB3, TE33, and 39-A11 with bound steroid (green), peptide (blue), and hapten 4 (purple), respectively.  $\text{Trp}^{H50}$ ,  $\text{Asn}^{H35}$ , and  $\text{Trp}/\text{Arg}^{H100}$  are also shown.

served length of CDRH3 and the conserved residue in position H100 result from antigen-driven selection during recombination. The length of the CDRH3 loop of IE9 is also conserved, and this antibody contains arginine at position H100. A similar combining site is used by TE33 to bind its peptide antigen. Four NH<sub>2</sub>-terminal residues of the peptide form a critical portion of an antiparallel  $\beta$  turn that is buried in the binding pocket. These residues are oriented in the same manner as are hapten 4 and progesterone in their respective binding pockets (Fig. 4). Again, the indole side chain of Trp<sup>H50</sup> plays a prominent role by packing against Gly<sup>5</sup> of the peptide. However, TE33 has a shorter CDRH3 loop (only seven amino acids), allowing the COOH-terminus of the peptide to exit the binding pocket.

Structural analysis of the esterolytic antibody 48G7 and its germ-line precursor suggested that, in addition to sequence diversity, conformational diversity intrinsic to some germline antibodies may contribute to the ability of the germ-line repertoire to bind such a wide array of chemical structures (2). The above structural analysis suggests that another important mechanism may involve the selection through evolution of a set of germ-line antibodies that are polyspecific—a feature shared by the major histocompatibility complex molecules, which bind many diverse peptides in the cellular immune response (28). Germ-line gene duplication may have created sets of closely related genes (V <sub>$\kappa$</sub> 1 and VGAM3.8) whose products contain subtle, but mechanically important, amino acid differences that, in conjunction with CDRH3 loop diversity, may allow presentation of several variations of the combining site. This combining site may have been selected during evolution as an optimal start point for rapid evolution of high-affinity, specific combining sites for a broad range of structures through a limited number of somatic mutations (these mutations may also remove interactions with idiotypic antibodies that regulate self recognition). Thus, the immune system likely relies on a variety of strategies, including conformational diversity and polyspecificity, in addition to somatic processes, to solve the problem of molecular recognition.

## REFERENCES AND NOTES

- P. G. Schultz and R. A. Lerner, *Science* **269**, 1835 (1995).
- P. A. Patten *et al.*, *ibid.* **271**, 1086 (1996); G. J. Wedemayer, P. A. Patten, L. H. Wang, P. G. Schultz, R. C. Stevens, *ibid.* **276**, 1665 (1997).
- H. Ulrich, *et al.*, *Nature* **389**, 271 (1997).
- G. W. Zhou, J. Guo, W. Huang, R. J. Fletterick, T. S. Scanlan, *Science*, **265**, 1059 (1994); L. C. Hsieh-Wilson, P. G. Schultz, R. C. Stevens, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5363 (1996); J. B. Charbonnier *et al.*, *Science* **275**, 1140 (1997).
- M. R. Haynes, E. A. Stura, D. Hilvert, I. A. Wilson, *Science* **263**, 646 (1994).
- A. Braisted and P. G. Schultz, *J. Am. Chem. Soc.* **112**, 7430 (1990).
- V. E. Gouverneur *et al.*, *Science* **262**, 204 (1993).
- D. Hilvert, K. W. Hill, K. D. Nared, M.-T. M. Auditor, *J. Am. Chem. Soc.* **111**, 9261 (1989).
- The variable regions of murine antibody 39-A11 and its germ-line precursor were cloned and expressed as humanized Fab fragments in the plasmid p4XH by fusing the V<sub>H</sub> (heavy) and V<sub>L</sub> (light) variable region genes to human C <sub>$\mu$</sub> 1 and C <sub>$\kappa$</sub>  (kappa) constant regions, respectively (29). This process introduces the following mutations into the 5' end of each variable region: Asp<sup>L1</sup>→Glu, Val<sup>H11</sup>→Leu, Ile<sup>H12</sup>→Val, Val<sup>H15</sup>→Leu, Gln<sup>H16</sup>→Glu. The Fab fragment was isolated from *E. coli* 25F2 transformed with the appropriate plasmid, as previously described (29). Fab concentrations of ~1.3 and 75 mg/liter were achieved in shaker flasks and a high-density fermentor, respectively. The Fab fragment was purified by protein G affinity chromatography (Sephacrose CL-6B, Pierce), followed by anion exchange chromatography (HS resin, Boehringer Mannheim) and passage through a gel filtration column (Pharmacia Superdex 200) immediately before crystallization trials. Mutant proteins were prepared by polymerase chain reaction (PCR)-mediated mutagenesis, and mutations were verified by sequencing the entire variable region gene.
- M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1678 (1977); T. C. Bruice and V. K. Pandit, *J. Am. Chem. Soc.* **82**, 5858 (1960); D. A. Storm and D. E. Koshland Jr., *Proc. Natl. Acad. Sci. U.S.A.* **66**, 445 (1970).
- A. D. Mesecar, B. L. Stoddard, D. E. Koshland Jr., *Science* **277**, 202 (1997).
- I. Fleming, *Frontier Orbitals in Organic Chemical Reactions* (Wiley, New York, 1976).
- Y. M. Chook, H. Ke, W. N. Lipscomb, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8600 (1993).
- Y. Xue, W. N. Lipscomb, R. Graf, G. Schnappauf, G. Braus, *ibid.* **91**, 10814 (1994).
- Y. Xue and W. N. Lipscomb, *ibid.* **92**, 10595 (1995).
- D. R. Liu, S. T. Cload, R. M. Pastor, P. G. Schultz, *J. Am. Chem. Soc.* **118**, 1789 (1996).
- H. Kim *et al.*, *J. Immunol.* **143**, 638 (1989).
- M. J. Sims, U. Krawinkel, M. J. Taussig, *ibid.* **149**, 1642 (1992).
- H. Gu, D. Kitamura, K. Rajewsky, *Cell* **65**, 47 (1991).
- E. A. Kabat, T. T. Wu, H. M. Perry, K. Gottesman, C. Foeller, *Sequences of Proteins of Immunological Interest* (Department of Health and Human Services, National Institutes of Health, Bethesda, MD, ed. 5, 1991).
- Antibody affinities for haptens were determined by fluorescence quenching [K. Taira and S. J. Benkovic, *J. Med. Chem.* **31**, 129 (1988)] with a Perkin-Elmer F-4500 luminescence spectrophotometer at 25°C in 20 mM Hepes-NaOH (pH 7.5) and 100 mM NaCl, with excitation at 280 nm and detection at 345 nm. Each K<sub>d</sub> value is the result of 15 independent titrations of antibody with hapten at two protein concentrations (with satisfactory agreement between the two experiments); K<sub>d</sub> values for the high- and low-affinity antibodies were determined at antibody concentrations of 50 and 250 nM, respectively. Reaction velocities were determined as described (6) at 25°C in a solution containing 20 mM Hepes-NaOH (pH 7.5), 100 mM NaCl, and 1% acetonitrile. The concentrations of the 39-A11 affinity-matured and germ-line antibodies were 1 and 3.6  $\mu$ M, respectively. Substrates were synthesized as described previously (6). The concentration of substrate was varied between 50 and 350  $\mu$ M, and all data were obtained in triplicate, with each experiment at a fixed substrate concentration also being performed in triplicate.
- J. H. Avevalo, E. A. Stura, M. J. Taussig, I. A. Wilson, *J. Mol. Biol.* **231**, 103 (1993).
- M. Shoham, *ibid.* **232**, 1169 (1993); T. Scherf, R. Hilles, F. Naider, M. Levitt, J. Anglister, *Biochemistry* **31**, 6884 (1992).
- M. Haynes *et al.*, *Israel J. Chem.* **136**, 151 (1996).
- E. Deverson, C. Berek, M. Taussig, A. Feinstein, *Eur. J. Immunol.* **17**, 9 (1987).
- J. H. Avevalo, M. J. Taussig, I. A. Wilson, *Nature* **365**, 859 (1993); J. H. Avevalo *et al.*, *J. Mol. Biol.* **241**, 663 (1994).
- R. Levy, O. Assulin, T. Scherf, M. Levitt, J. Anglister, *Biochemistry* **28**, 7168 (1989).
- H.-G. Rammensee, K. Falk, O. Rötzschke, *Annu. Rev. Immunol.* **11**, 213 (1993); J. B. Rothbard and M. L. Gefter, *ibid.* **9**, 527 (1991).
- H. D. Ulrich *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11907 (1995).
- J. Navaza, *Acta Crystallogr. A* **50**, 157 (1994).
- The incomplete factorial crystallization screen [J. Jancarik and S. H. Kim, *J. Appl. Crystallogr.* **24**, 409 (1991)] was used to identify the following crystallization conditions for 39-A11: 100 mM tris (pH 8.0), 17% polyethylene glycol (PEG) 8000, 200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 10 mM CdSO<sub>4</sub>. Crystals (0.6 mm by 0.6 mm by 0.2 mm) were grown by hanging drop vapor diffusion at 4°C from 2  $\mu$ l of mother liquor and 2  $\mu$ l of protein solution (12 mg/ml) with stoichiometric hapten in 10 mM tris (pH 8.0), 100 mM NaCl, and 1 mM methionine. The crystals (39-A11) were frozen in liquid nitrogen after a brief washing (10 s) in 15% PEG 8000 and 25% glycerol mixed 1:1 with mother liquor. Crystals of the germ-line antibody were grown similarly (without hapten) from 19% PEG, 20% glycerol, 200 mM magnesium acetate, 100 mM bis-tris (pH 6.0), and 10 mM CdSO<sub>4</sub>, and were frozen with 30% glycerol. This approach allowed complete data sets to be collected from single crystals at -165°C on Stanford Synchrotron Radiation Laboratory beamlines 7-1 (mature) and 9-1 (germ line). The reflections were indexed (DENZO) and scaled and merged with SCALEPACK (30). The structure of the mature Fab was solved by molecular replacement with the program package AMORE (30), with coordinates from 1HKL and 1TET (Brookhaven protein database) as search models for the constant and variable regions, respectively [M. Shoham, P. Proctor, D. Hughes, E. T. Baldwin, *Proteins* **11**, 218 (1991)]. The crystal form has two molecules per asymmetric unit. Initially, the first molecule was found and the elbow angle refined by PC refinement. The second molecule was located with the fix command in AMORE and searching with the PC-refined solution. The molecule was then refined with X-PLOR version 3.851 [A. Brünger, *X-PLOR, Version 3.851. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992)] PC refinement of the separate domains, followed by simulated annealing and alternating rounds of model building with O [T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. A* **47**, 110 (1991)] and positional refinement. The final structure of 39-A11 contains 869 amino acids, 2 hapten molecules, 213 water molecules, and 4 Cd<sup>2+</sup> ions. The germ-line Fab crystallized in the space group C2 with two molecules in the asymmetric unit. The structure was solved similarly with the use of the mature structure as a search model; the final structure contains 868 amino acids and 274 water molecules.
- Enzyme-linked immunosorbent assay plates were coated with 100  $\mu$ l of hapten-BSA conjugate (10  $\mu$ g/ml) in phosphate-buffered saline and incubated overnight at 4°C. Plates were washed, blocked (with phosphate-buffered saline containing 1% BSA and 0.05% Tween 20), and washed again. They were then incubated with 100  $\mu$ l of 100 nM antibody, washed, and incubated with alkaline phosphatase-conjugated antibodies to human kappa chain. After washing *p*-nitrophenylphosphate was added to each well and absorbance was measured at 405 nm.
- Supported by the U.S. Department of Energy, NIH, Howard Hughes Medical Institute (P.G.S.), and W. M. Keck Foundation. We thank Stanford Synchrotron Radiation Laboratory for data collection time. The coordinates have been deposited in the Brookhaven protein database under accession numbers 1A4K (mature) and 1A4J (germ line).

15 September 1997; accepted 12 January 1998