Using antibody catalysis to study the outcome of multiple evolutionary trials of a chemical task

Amelie Karlstrom, Guofu Zhong, Christoph Rader, Nicholas A. Larsen, Andreas Heine, Roberta Fuller, Benjamin List, Fujie Tanaka, Ian A. Wilson, Carlos F. Barbas III*, and Richard A. Lerner*

The Skaggs Institute for Chemical Biology and the Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037

Contributed by Richard A. Lerner, December 27, 1999

Catalytic aldolase antibodies generated by immunization with two different, but structurally related, β-diketone haptens were cloned and sequenced to study similarities and differences between independently evolved catalysts. Kinetic and sequence analysis coupled with mutagenesis, structural, and modeling studies reveal that the defining event in the evolution of these catalysts was a somatic mutation that placed a lysine residue in a deep, yet otherwise unrefined, hydrophobic pocket. We suggest that covalent chemistries may be as readily selected from the immune repertoire as the traditional noncovalent interactions that have formed the basis of immunochemistry until this time. Further, we believe that these experiments recapitulate the defining events in the evolution of nature's enzymes, particularly as they relate to chemical mechanism, catalytic promiscuity, and gene duplication.

One of the most fascinating questions in biology concerns what would happen if evolution started over again. For any system within an organism, the answer must, in part, depend on how many ways there are to accomplish a given task. For example, in the case of protein enzymes, the problem reduces to the purely chemical issue of the potential of the amino acids acting in concert to catalyze a given transformation. If there are many possible routes, multiple solutions to the problem may occur, such as for hydrolysis of the amide bond. However, even where there are multiple solutions, reoccurring motifs are found such as the catalytic triad of the serine proteases. The study of these questions is experimentally difficult because, other than "gedanken" experiments, it is generally not possible to restart the evolution of a system. However, the advent of the process of reactive immunization to generate antibody catalysts allows examination of the outcome of multiple evolutionary trials to accomplish a complicated chemical task by different members of a protein family.

Reactive immunization differs from the more usual immunization procedure in that reactive chemicals are used as immunogens. Thus, the selection parameters of the immune system are switched from simple binding to more complicated chemistry, such as the formation of a covalent bond between the antibody and the antigen. When the selectable chemistry is part of a catalytic mechanism, the result is often evolution of an efficient enzyme in real time (1–5).

Recently, we have used the reactive immunization procedure to generate catalytic antibody aldolases that are remarkably similar in efficiency and mechanism to the natural aldolases (1–3). One essential feature of the antibody aldolases is an active site containing a deeply buried lysine residue with a highly perturbed pK_a was selected. This reactive lysine interacts with aldol donors to form an enamine that is the nascent carbon nucleophile that adds into the aldol acceptor to form the new carbon-carbon bond yielding a β-hydroxy ketone product with precise control of the stereochemistry at the carbinol carbon. Thus, reactive immunization allowed selection for a complex chemical process that requires, among other parameters: (*i*) selection of a binding pocket of suitable configuration to interact with two substrates in a bimolecular process; (*ii*) an active site with an accessible lysine residue; (*iii*) perturbation of the lysine pK_a; (*iv*) the presence of a base; and (*v*) orchestration of the interaction with water allowing for the formation and hydrolysis of imine intermediates.

In the present work, we challenged the immune system with the same chemical task to study the outcome of multiple evolutionary attempts to create an enzyme. Here, we present our studies of antibody aldolases that exhibit the same enantioselectivity, the same apparent chemical mechanism, but were derived by independent immunization of mice with diketone haptens **1** and **2** (Fig. 1).

Materials and Methods

Cloning of Antibody Genes. The antibody genes from hybridoma cell lines expressing the catalytic antibodies 38C2, 33F12, 40F12, and 42F1 were cloned as described (6, 7).

Site-Directed Mutagenesis. Lysine 93 in the heavy chain of 33F12 and 40F12 was mutated to an alanine residue by PCR using two overlapping oligonucleotides carrying the desired substitution. PCR was carried out by using the proof-reading Expand High Fidelity PCR System *Pow*/*Taq* DNA polymerase mix (Roche Molecular Biochemicals) to minimize PCR errors. The correct sequences of the mutants were confirmed by DNA sequencing.

Subcloning into the pIGG Expression Vector. The pIGG expression vector used here is the expression of whole human IgG in mammalian cells (C.R. and C.F.B., unpublished work). The heavy chain variable domains of the antibodies were PCR-amplified by using Expand High Fidelity PCR System *Pow*/*Taq* DNA polymerase mix and antibody-specific primers incorporating the restriction sites required for pIGG cloning. The PCR products were digested with *Sac*I and *Apa*I and were ligated to *Sac*I/*Apa*I-digested pIGG vector by using T4 DNA ligase (Life Technologies, Gaithersburg, MD). The ligation reactions were transformed into *Escherichia coli* strain XL-1 Blue (Stratagene) by electroporation. The light chain variable domains of the antibodies were PCR-amplified and were fused to a human kappa chain by overlap extension PCR using Expand High Fidelity PCR System *Pow*/*Taq* DNA polymerase mix and antibody-specific primers incorporating the restriction sites required for pIGG cloning. The PCR products were digested with *Spe*I and *Xba*I and were ligated to the vector with the appropriate heavy chain variable domain, prepared by digestion with *Spe*I and *Xba*I and 5′-desphosphorylation with alkaline phosphatase from calf intestine (Roche Molecular Biochemicals). The ligation reactions were transformed into XL-1 Blue by electroporation. The correct sequences of the pIGG-antibody clones were confirmed by DNA sequencing.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF242212–AF242219).

*To whom reprint requests should be addressed. E-mail: foleyral@scripps.edu or carlos@scripps.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

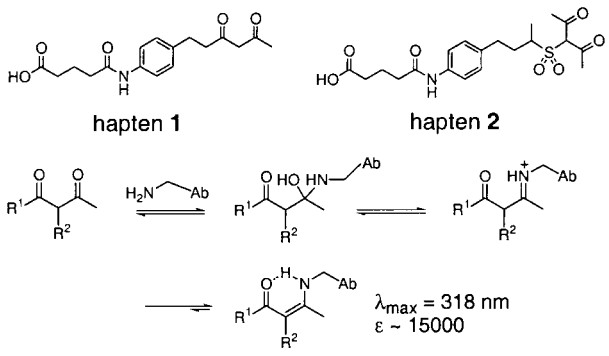


Fig. 1. Aldolase antibodies were generated by reactive immunization with 1,3-diketone hapten **1** or **2.** A stable covalent interaction is formed on reaction of the haptens with the aldolase antibodies. The multistep reaction yields a stable vinylogous amide or enaminone. Unlike hapten **1**, hapten **2** incorporates a sulfone group to mimic the tetrahedral geometry of the rate-determining transition state of carbon-carbon bond formation.

Expression and Purification of Recombinant Antibodies. The recombinant hybrid mouse/human IgG antibodies were expressed in transiently transfected human embryonic kidney 293 T cells. Near-confluent cells were transfected with pIGG-antibody plasmid DNA by LipofectAMINE PLUS (Life Technologies) according to the procedure recommended by the manufacturer. After transfection, the cells were cultured for 3 days in DMEM (Life Technologies) supplemented with 10% FBS (Life Technologies). Cell culture supernatant was collected and replaced with fresh medium every day. IgG were purified from concentrated cell culture supernatants by affinity chromatography using a 1-ml Protein A HiTrap column (Pharmacia) attached to an FPLC system (Pharmacia). PBS was used as equilibration and washing buffer and 0.5 M acetic acid for elution. Eluted fractions were neutralized immediately by using 0.5 volumes of 1 M Tris·HCl (pH 9.0) and were pooled, concentrated, and exchanged into PBS.

Activity Assays. Aldolase activity of the recombinant antibodies was measured by using the fluorogenic *retro*-aldol substrate methodol (8). The reactions were carried out in PBS (pH 7.4), with the substrate diluted from a 5 mM stock solution in acetonitrile to a final concentration of 200 μM. The assays were performed at room temperature in microtiter plates in a reaction volume of 100 μl. An antibody concentration of 1 μM was used. Formation of the aldehyde product was followed with a fluorescence plate reader by monitoring at λ_{ex} = 330 nm and λ_{em} = 452 nm.

Computer-Based Homology Model. The template for the homology model of catalytic antibody 40F12 was selected by sequence comparison to our database of 120 structurally determined Fabs. Specifically, we used the V_L region of Fab MOPC21 (9) and the V_H region of Fab 33F12 (2), which had the greatest sequence identity to the V_L and V_H, respectively, of 40F12. The chosen V_L and V_H templates also had matching CDR lengths, with the exception of the heavy chain CDR3 loop (H3). Inspection of the database showed that Fab 730.1.4 (10) had an H3 loop with the same length. A least squares fitting of the Cα traces of 33F12 and 730.1.4, 10 residues before and 10 residues after the H3 loop, showed residues H92 and H104 were the points from which the CDR3 structures diverged. Hence, the H3 loop of 730.1.4 was spliced into 33F12 between residues H92 and H104. Where necessary, residues in this preliminary model were mutated to conform to the 40F12 sequence, and the conformations of conflicting sidechains were altered based on the standard rotamer library (11). Finally, the model was refined in two rounds of energy minimization with the program DISCOVER (Biosym Technologies, San Diego). In the first round, the mutated sidechains were minimized with the rest of the model constrained, and in the second round all sidechains were minimized with the backbone constrained.

Results and Discussion

A Second Evolutionary Trial. Our second evolutionary trial began with the synthesis of hapten **2** and the generation of monoclonal antibodies (5). Of 11 antibodies that were shown to catalyze the aldol reaction, antibodies 40F12 and 42F1 catalyzed aldol addition reactions and *retro*-aldol reactions with enantioselectivities similar to that previously observed for antibodies 33F12 and 38C2 (1–3). The remaining nine catalytic antibodies exhibited a reversal in enantioselectivity, preferentially forming aldols antipodal to those formed with 33F12 or 38C2. These antipodal catalysts are the subject of another report (5). We have restricted this analysis of aldolase antibodies to those catalysts that exhibit the same overall enantioselectivity in their catalysis of aldol reactions.

Antibodies 40F12 and 42F1 were first identified by their ability to react covalently with 2,4-pentanedione to form a stable enaminone (UV at λ_{max} 318 nm) (1). Like antibodies 33F12 and 38C2, 40F12 and 42F1 showed the characteristic enaminone absorption maximum at 318 nm after incubation with 2,4-pentanedione (Fig. 2*A*). The antibodies were then studied for their ability to catalyze a variety of aldol and *retro*-aldol reactions (Table 1). All antibody catalyzed aldol and *retro*-aldol reactions followed Michaelis-Menten kinetics and were inhibited by addition of a stoichiometric amount of 2,4-pentanedione. A summary of the kinetic parameters for antibody-catalyzed aldol and *retro*-aldol reactions is provided (Table 1). These results are consistent with the programming of a reactive amine in the covalent catalytic mechanism (1–3). Further, these antibodies recapitulate the broad catalytic scope that was first observed with aldolase antibodies 33F12 and 38C2.

Sequence Analysis. To study the aldolase antibodies at the level of gene and protein sequence, their cDNAs were cloned, and expressed proteins were characterized. Alignment of aldolase antibody protein sequences shows that the two antibodies 38C2 and 33F12, generated by immunization with hapten **1**, are highly homologous, contain the same V_H CDR3 sequence and, therefore, originate from the same germline antibody. They differ from one another by 10 amino acid substitutions in V_L and 9 in V_H as a consequence of somatic mutation (Fig. 3) (2). Similarly, the amino acid sequences of the two antibodies 40F12 and 42F1, generated by immunization with hapten **2**, are highly homologous and differ from each other only in 3 positions in V_L and in 2 positions in V_H. They too are somatic variants of one another. When the sequences of these two sets of antibodies are compared, it is evident that all of the antibodies have high homology in their V_H gene segments. Searches of the available databases, including the IMGT database (13), suggest that the heavy chains of both sets of antibodies could be derived from the same germline gene because their nucleic acid sequences were found to be most homologous to the V_H 22.1 segment (Fig. 3). Because the database of the murine germline repertoire is incomplete, we cannot, without additional cloning, exclude the existence of another more homologous V_H gene. Although their component V_H gene segments are highly homologous, the two sets of antibodies use very different V_H CDR3 regions for antigen recognition. The 40F12/42F1 family uses an H3 segment that is five amino acids longer that that used by the 33F12/38C2 family. In contrast to the V_H segment homology, the light chain V_L sequences of the two sets of antibodies are highly divergent and are derived from different germline genes.

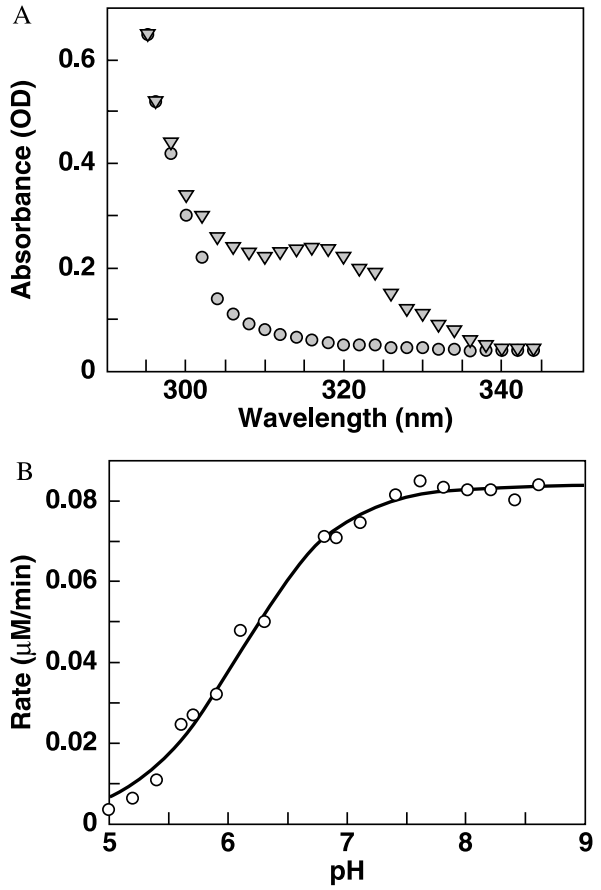


Fig. 2. (*A*) Aldolase antibody 40F12 (75 μM binding sites) with 4 eq. of 2,4-pentanedione presented the typical absorption maximum of the enaminone at 318 nm (shown in orange) whereas inactive antibodies did not (shown in blue). (*B*) Rate of enaminone formation as a function of pH. Enaminone formation between antibody 40F12 and 2,4-pentanedione was followed spectrophotometrically at 318 nm at 22°C. The incubation mixtures contained 15 μM antibody binding sites and 50 μM 2,4-pentanedione in citrate-phosphate buffer in the pH range from 5.0 to 8.6. The reaction rates were calculated with the use of the experimentally determined extinction coefficient ε_{318} = 17.4 $mM^{-1}{\cdot}cm^{-1}$.

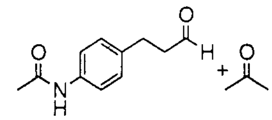
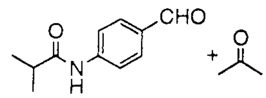
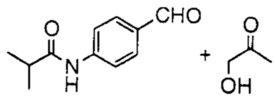
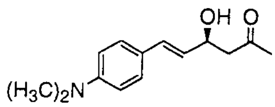
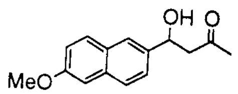
A Conserved and Chemically Reactive Lysine. Previous studies of antibodies 33F12 and 38C2 showed that these antibodies use a catalytic mechanism analogous to natural class I aldolase enzymes (1–3). In this mechanism, the ε-amino group of a lysine residue located in the active site acts as a nucleophile to form an enamine with a donor ketone substrate, and subsequent reaction with an acceptor aldehyde substrate provides, after hydrolysis of the imine, the aldol product. The crystal structure of 33F12 (2) revealed that Lys 93 of the heavy chain [numbering according to Kabat (13)] contributes the best candidate catalytic ε-amino group to the antibody binding pocket. Strikingly, lysine H93 is conserved in all four antibody sequences, although this residue is most typically conserved as alanine. Lysine at position H93 is very unusual, occurring in just 11 known antibody sequences (14). Comparison with the putative germline residue at this position indicates that the occurrence of lysine is the result of a somatic mutation, as no known germline V_H segment encodes lysine at this position.

For the ε-amino group of Lys to be nucleophilic, it must be uncharged. However, the pK_a of the ε-amino group in the free amino acid lysine in aqueous solution is 10.5 (15). Because catalytic activity depends on a nonprotonated lysine as a nucleophile and aldolases operate with maximal activity at neutral pH where the ε-amino group of lysine normally would be protonated, the pK_a of this group must be perturbed. To approximate the pK_a of the essential lysine, we have used a kinetic assay based on the ability of the antibodies to form enamines with β-diketones. We studied the reaction of 2,4-pentanedione with antibody 40F12 and spectrophotometrically monitored the absorption of the antibody-enamine complex at 318 nm. The pH dependence of this reaction is shown and is described by a simple titration curve with a pK_a of 6.0 for 40F12 (Fig. 2*B*). Previous studies of antibodies 33F12 and 38C2 revealed pK_as of 5.5 and 6.0, respectively, for their reactive lysines (2). The dependence of k_{cat}/K_M as a function of pH for the *retro*-aldol reaction also showed an acidic limb pK_a = 6.3–6.9 for the family of aldol antibodies. This result approximates the pH dependence of catalysis of FDP aldolase (16). Thus, the aldol antibodies exhibit a chemically reactive amine group with a highly perturbed pK_a in their active sites.

Mutational Studies. To confirm the identity of the chemically reactive lysine at position H93, site-directed mutagenesis studies of antibodies 33F12 and 40F12 were performed. Two mutants were constructed in which Lys^{H93} was substituted with Ala. For efficient expression of the antibodies, the variable domains were cloned into a mammalian expression vector in which the heavy and the light chains were expressed from two CMV promoters to give an IgG_1 hybrid antibody with a human kappa constant region and human γ1 constant region (C.R. and C.F.B., unpublished data). The antibodies were expressed in a human cell line and were secreted into the cell culture medium, from which they were purified by protein A affinity chromatography. Subsequent analysis of the catalytic activity of the mutants and the wild-type hybrid antibodies was carried out by using the aldol sensor methodol, which, on catalytic conversion, provides a fluorescent product. This *retro*-aldol reaction was used because its fluorescent readout is the most sensitive indicator of aldolase catalytic activity (8). Analysis of recombinant 33F12 and 40F12 indicated that they maintained the catalytic activity of the hybridoma expressed antibody. However, the corresponding $Lys^{H93}{\rightarrow}$Ala mutants of these antibodies showed a complete loss of catalytic activity. These results demonstrate the functional significance of this residue in both antibody families. Although structural alterations are difficult to exclude, the substitution of lysine with alanine is most likely tolerated at this position because alanine is the amino acid most typically found at this position. ELISA binding studies of the $Lys^{H93}{\rightarrow}$Ala mutants to haptens **1** and **2** demonstrated that their specific binding to either of these diketones was abrogated. All of the hybridoma-expressed and recombinant aldolase antibodies bound both haptens with high affinity. Thus, these result imply that the loss of activity in the $Lys^{H93}{\rightarrow}$Ala mutants is most likely caused by the substitution of a residue essential for the engagement of diketones in a covalent enaminone linkage, consistent with the notion that this residue is the key catalytic lysine.

Homology in the Active Sites of Aldolase Antibodies and a Natural Aldolase Enzyme. In previous studies, we have reported the crystal structure of antibody 33F12 and an analysis of its active site. To gain insight into the active site of the 40F12 antibody, we constructed a model of the antibody by sequence comparison of 40F12 to our database of known Fab structures. The crystal structure of 33F12 and the homology model of 40F12 both show the reactive lysine at the bottom of a deep, hydrophobic binding pocket (Fig. 4 *A* and *B*) (2). In 33F12 and 40F12, most of the residues within a 5-Å radius of the ε-amino group of lysine H93 are hydrophobic (Table 2). Further, the vast majority of these residues are encoded within the germline gene segment V_H 22.1, with only His^{H35} in antibodies 40F12 and 42F1 resulting from somatic hypermutation. The dominant role of the V_H segment in

Table 1. Kinetic parameters for antibody-catalyzed aldol and *retro*-aldol reactions

Substrate	Antibody	k_{cat}, min^{-1}*	K_m, μM*	k_{cat}/k_{un}†	$(k_{cat}/K_m)/k_{un}$‡
	40F12	0.0054	33	2.4×10^4	7.2×10^8
	42F1	0.0026	41	1.1×10^4	2.8×10^8
	38C2§	0.0067	17	2.9×10^4	1.7×10^9
	40F12	0.071	203	6.5×10^5	3.2×10^9
	42F1	0.048	271	4.3×10^5	1.6×10^9
	38C2	0.21	123	1.9×10^6	1.5×10^{10}
	40F12	0.21	317	8.8×10^5	2.8×10^9
	42F1	0.14	266	5.8×10^5	2.2×10^9
	38C2	1.1	184	4.6×10^6	2.5×10^{10}
	40F12	0.11	130	6.1×10^2	4.7×10^6
	33F12	0.21	317	1.2×10^3	3.8×10^6
	38C2	4.8	20	2.7×10^4	1.4×10^9
	40F12	0.11	130	1.1×10^5	8.5×10^8
	33F12	0.11	43	1.1×10^5	2.6×10^9
	38C2¶	1.0	14	1.0×10^6	7.1×10^{10}

*The kinetic data of k_{cat} and K_m (per antibody active site) were obtained in PBS at pH 7.4 by fitting experimental data to nonlinear regression analysis using GRAFIT software.
†Aldol reactions with a unit (M).
‡*Retro*-aldol reactions with a unit (M^{-1}).
§Refs. 1 and 2.
¶Ref. 8.

shaping the binding sites of these antibodies around the critical lysine accounts for the ability of these two families of antibodies to accomplish the same catalytic task while using very different light chains. The similar architecture of the antibody binding sites in the two families of antibodies suggests that these aldolases function through analogous catalytic mechanisms using Lys^{H93} as the nucleophile. The pK_a of the ε-amino group of this lysine residue is, therefore, most likely perturbed by the hydrophobic environment of the binding sites present in both antibody families, as we originally proposed for antibody 33F12 (2). This hydrophobic microenvironment disfavors protonation and charge development on the ε-amino group of Lys^{H93}, thereby reducing its pK_a. To study the relationship of our antibody aldolases with their natural enzyme counterparts, we surveyed the crystal structures of natural aldolase enzymes. Analysis of the x-ray structure of *N*-acetylneuraminate lyase revealed a deep pocket, which is predicted to be the active site of the enzyme (17). In this enzyme, Lys165 lies at the bottom of the active site pocket in a hydrophobic environment (Fig. 4*C*) and has been predicted to be the key lysine residue in the catalytic mechanism

V_L amino acid sequences

```
                 FR1                    CDR1              FR2          CDR2              FR3                      CDR3        FR4
33F12  DVVMTQTPLSLPVSLGDQASISC  RSSQSLVHSYGNTFLN  WYLQKSGQSPKLLIY  KVSNRFS  GVPDRFSGSGSGTDFTLKISRVEAEDLGVYFC  SQGTHVPYT  FGGGTKLELK
38C2   -------------R---------  ------L-T--SPY--  -----P---------  -------  -----------------R--------------  -----L---  --------I-

40F12  DIKMTQSQKFMSTSVGDRVTVTC  KASQNVGRNVG       WYQQKPGQSPKSLIY  STSYRYS  GVPDRFTGSGSGTDFTLTVNNVQSEDLADYFC  HQYNSYPFT  FGSGTKLEIK
42F1   --V--------------------  -------K---       ---------------  -------  --------------------------------  ---------  --------L-
```

V_H amino acid sequences

```
                       FR1                CDR1       FR2                CDR2                         FR3                      CDR3          FR4
GERM   EVKLEESGGGLVQPGGSMKLSCVASGFTFS  NYWMS  WVRQSPEKGLEWVA  QIRLKSDNYATHYAESVKG  RFTISRDDSKSSVYLQMNNLRAEDTGIYYCTG
33F12  ----V---------E--------V--L---  RF---  --------------  E------------------  K----------RL-----S--T--------KI  YFYSFSY      WGQGTLVTVS
38C2   ----V-----------T-----EI--L--R  -----  --------------  E---R--------------  K----------RL-----S--T--------KT  -------      ----------
40F12  --Q-QQ------------------------  Y---H  --------------  D-----ND--S--------  ------------------------------KG  ASLQRLLYTTAY  -----S----
42F1   --Q-QQ------------------------  Y---H  --------------  D-----N---S--------  ---------------------P--------KG  ASLQRLLYTTAY  -----S----
```

Fig. 3. Amino acid sequence alignment of aldolase antibodies. The framework regions (FR) and complementarity determining regions (CDR) are indicated. Dashes indicate identical amino acids. In contrast to the V_L sequences, the V_H sequences of the four aldolase antibodies are highly related. Databank screening revealed V_H 22.1 as the closest related germline gene.

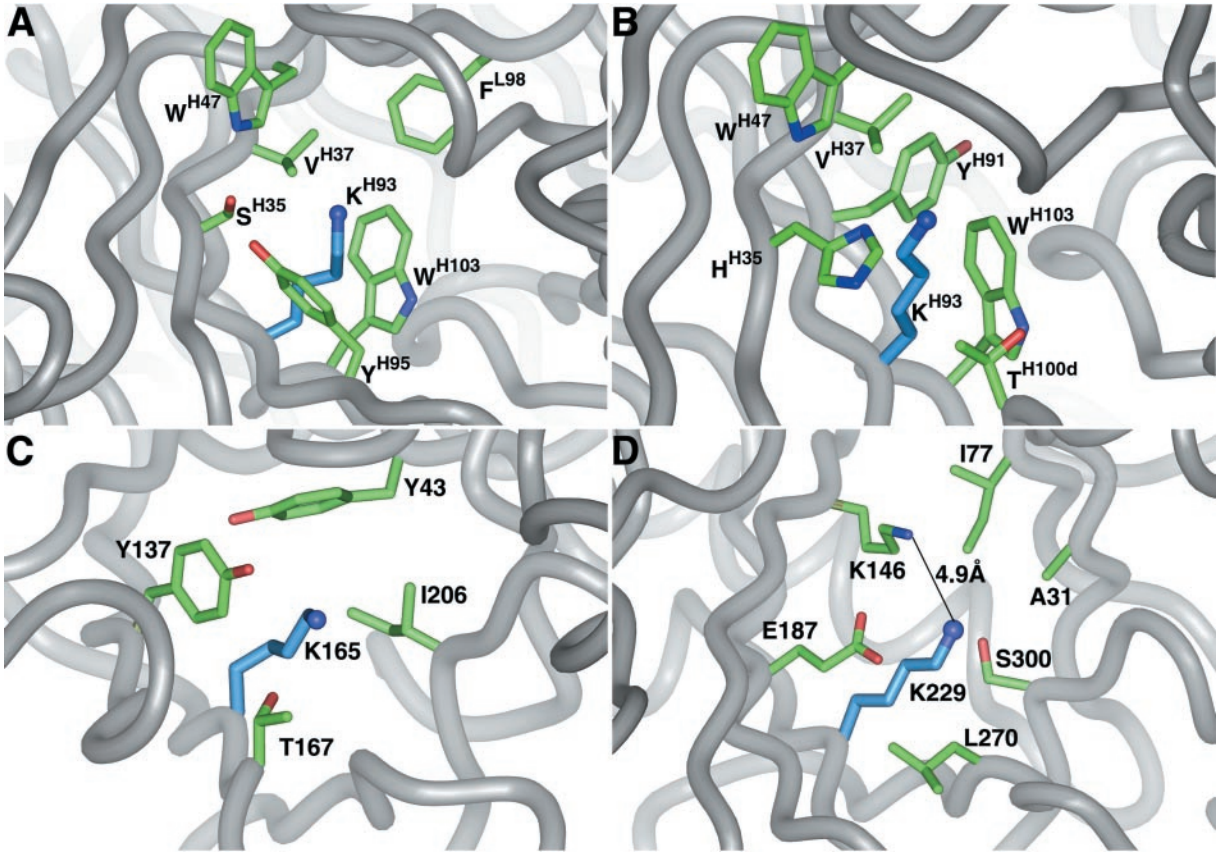


Fig. 4. Comparison of aldolase abzyme and enzyme active sites. All of the labeled side chains are within 5 Å of the catalytic lysine (blue) Nε group. (*A*) Top view of the Fab′ 33F12 combining site. (*B*) Same view for the homology model of 40F12. (*C*) View of the *N*-acetylneuraminate lyase active site (17) around the Nε of the catalytic lysine residue 165. (*D*) Same view of fructose 1,6- bisphosphate aldolase active site (19) around the Nε of the catalytic lysine residue 229. The figure was generated with INSIGHT (Biosym Technologies).

of this aldolase. As with our aldolase antibodies, a hydrophobic microenvironment may be responsible for tuning the pK_a of the ε-amino group of this lysine residue. The majority of natural aldolase enzymes appear, however, to use an electrostatic mechanism for the perturbation of the pK_a of the ε-amino group of the lysine. This mechanism was first proposed by Westheimer in his studies of another amine-dependent enzyme, acetoacetate decarboxylase (ref. 18 and references therein). Multiple lysine residues are observed in the active sites of a variety of class I fructose 1,6-bisphosphate aldolases, suggesting that electrostatics might account for increased reactivity of the key lysine residue in this common glycolytic enzyme. In rabbit muscle aldolase (19), the distance between the ε-amino groups of Lys229 and Lys146 is 4.9 Å. (Fig. 4*D*), comparable to that observed for mandelate racemase (20) in which electrostatic effects are also believed to be critical in the perturbation of the pK_a of a key ε-amino group. Comparison of the side chains present within a 5-Å radius of the active site amino group shows the radical differences in the types of side chains found at the active sites of enzymes that use either a hydrophobic or electrostatic mechanism for lysine pK_a perturbation.

Table 2. Residues within 5 Å of the Nε group of the catalytic lysine (shown in bold)

A	B	C	D
LysH93	**LysH93**	**Lys165**	**Lys229**
SerH35	HisH35	Tyr43	Ala31
ValH37	ValH37	Tyr137	Ile77
TrpH47	TrpH47	Thr167	Lys146
TyrH95	TyrH91	Gly189	Glu187
TrpH103	Thr H100d	Ile206	Leu270
PheL98	TrpH103		Ser300

A, aldolase catalytic antibody Fab′ 33F12 [PDB ID code 1AXT (2)]; B, homology model of aldolase antibody 40F12; C, molecule 4 of *N*-acetylneuraminate lyase [PDB ID code 1NAL (17)]; D, molecule D of fructose 1, 6-bisphosphate aldolase from rabbit muscle [PDB ID code 1ADO (19)].

Conclusions

The formation of antibodies is an evolutionary process that selects on binding energy. If one compares the way that new functions are generated in biology in general to that of the immune system, many similarities are observed. In each case, genotypic parameters that generate new information are coupled to a selection system. The main differences are in the selection parameters and the time line. General evolution uses natural selection to select for function over a long time whereas the immune system uses clonal selection to select for improvement in the binding energy of antibodies in a matter of weeks. Reactive immunization allows one to switch the selection parameters of the immune system to those more like natural selection while preserving the favorable time line of antibody induction. One interesting feature of these experiments is that, because of threshold considerations, there was no assurance that, even if the system were offered the possibility to accomplish complex chemistry, it would be a selectable parameter. For example, in the experiments reported here, the system was set up

so that it was possible for the system to drive toward formation of a covalent bond, which is the "ultimate" in achieving binding energy. However, if the threshold for cell division and differentiation were much lower, this rare event might never be seen. For example, if simple noncovalent binding to the benzene ring in the antigen gave sufficient binding energy for clones to be propagated, this feature becomes the selectable event with no driving force toward the more complicated chemistry required for catalysis of the aldol reaction. Nevertheless, a high percentage of clones achieve the ability to form a covalent bond; hence, a system evolved for simple binding can accomplish much more when the selection parameters are changed. Aside from the evolution of catalysis, the immune system must have a true evolutionary drive toward binding energy, and, when superior pathways toward this goal are offered, they may be taken.

Our studies provide a picture of the evolutionary events that lead to the catalytic mechanism and substrate promiscuity observed for this family of catalysts. In both evolutionary trials, a lysine residue appeared early during somatic refinement in germline antibodies using the putative V_H 22.1 germline segment. This germline appears to encode a deep hydrophobic pocket. The pocket in itself is not predisposed to bind the diketone haptens that we have studied here with high affinity, but a single somatic mutation at Lys^{H93} provides a covalent chemical solution for high affinity binding. The chemical reactivity that Lys^{H93} is endowed with in this type of binding site is readily selectable. Once this covalent chemistry appeared, further refinement of the binding pocket did not occur because it afforded no selective advantage to the system. This result complements studies of the ontogeny of other antibody catalysts and of other hapten binding antibodies in which noncovalent mechanisms are used (21–24). These studies have revealed that particular germline genes are poised for catalysis of particular reactions or for rapid evolution to high affinity hapten binding. Thus, we find a similar result when covalent chemistry is involved and suggest that covalent chemistries may be as readily selected from the immune repertoire as traditional noncovalent interactions that have shaped our thoughts regarding immunochemistry. Further, as we have discussed previously (2), we believe that these experiments recapitulate the defining events in the evolution of nature's enzymes, particularly as they relate to chemical mechanism, catalytic promiscuity, and gene duplication (25, 26).

Thus, one of the key results from this study is that some parameters for aldol catalysts are critical whereas others are more degenerate. Although the four aldolase antibodies described here share the same V_H germline segment, they exhibit very different V_H CDR3 and light chain sequences. Although the structural correlates for covalent mechanism within this class of antibodies are now defined, the structural correlates for catalytic efficiency have yet to be uncovered. Nevertheless, we already know that the catalytic proficiencies of antibody aldolases can exceed 10^{13} in some cases (5).

By necessity, our study depends on the chemical solutions allowed within an antibody framework. It is possible that alternative solutions might be favored in other protein frameworks. For example, we did not find a major solution to the lysine pK_a problem that nature most commonly uses. In most natural aldolases, a second lysine appears in the vicinity of the lysine nucleophile and perturbs its pK_a through electrostatic interactions. It might simply be that this second lysine so destabilizes the antibody fold that is it is never seen or that the germline repertoire is poised structurally favoring the hydrophobic mechanism, as it arises from the common use of the same germline V_H 22.1 segment by all four antibodies. Alternatively, the appearance of a second lysine may be a rarer event, and our sample size may have been too small to observe it.

This study was supported by Grant CA27489 (to I.A.W., C.F.B., and R.A.L.) from the National Institutes of Health and the Skaggs Institute for Chemical Biology.

1. Wagner, J., Lerner, R. A. & Barbas, C. F., III (1995) *Science* **270,** 1797–1800.
2. Barbas, C. F., III, Heine, A., Zhong, G., Hoffmann, T., Gramatikova, S., Björnestedt, R., List, B., Anderson, J., Stura, E. A., Wilson, I. A. & Lerner, R. A. (1997) *Science* **278,** 2085–2092.
3. Hoffmann, T., Zhong, G., List, B., Shabat, D., Anderson, J., Gramatikova, S., Lerner, R. A. & Barbas, C. F., III (1998) *J. Am. Chem. Soc.* **120,** 2768–2779.
4. Wirsching, P., Ashley, J. A., Lo, C. H., Janda, K. D. & Lerner, R. A. (1995) *Science* **270,** 1775–1782.
5. Zhong, G., Lerner, R. A. & Barbas, C. F., III (1999) *Angew. Chem. Int. Ed. Engl.* **38,** 3738–3741.
6. Rader, C., Cheresh, D. A. & Barbas, C. F., III (1998) *Proc. Natl. Acad. Sci. USA* **95,** 8910–8915.
7. Andris-Widhopf, J., Steinberger, P., Rader, C., Silverman, G. J., Fuller, R. P. & Barbas, C. F., III (2000) in *Phage Display: A Laboratory Manual*, eds. Barbas, C. F., III, Burton, D. R., Scott, J. K. & Silverman, G. J. (Cold Spring Harbor Lab. Press, Plainview, NY).
8. List, B., Barbas, C. F., III, & Lerner, R. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 15351–15355.
9. Derrick, J. P. & Wigley, D. B. (1994) *J. Mol. Biol.* **243,** 906–918.
10. Ban, N., Escobar, C., Garcia, R., Hasel, K., Day, J., Greenwood, A. & McPherson, A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1604–1608.
11. Ponder, J. W. & Richards, F. M. (1987) *J. Mol. Biol.* **193,** 775–791.
12. Lefranc, M. P., Giudicelli, V., Ginestoux, C., Bodmer, J., Muller, J., Bontrop, R., Lemaitre, M., Malik, A., Barbie, V. & Chaume, D. (1999) *Nucleic Acids Res.* **27,** 209–212.
13. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. & Foeller, C. (1991) *Sequences of Proteins of Immunological Interest* (U.S. Dept. of Health and Human Services, National Institutes of Health, Bethesda, MD).
14. Martin, A. C. R, (1996) *Proteins Struct. Funct. Genet.* **25,** 130–133.
15. Dean, J. A. (1992) in *Lange's Handbook of Chemistry* (McGraw–Hill, San Francisco, CA), p. 819.
16. Morris, A. J. & Tolan, D. R. (1994) *Biochemistry* **33,** 12291–12297.
17. Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. & Colman, P. M. (1994) *Structure (London)* **2,** 361–369.
18. Westheimer, F. H. (1995) *Tetrahedron* **51,** 3–20.
19. Blom, N. & Sygusch, J. (1997) *Nat. Struct. Biol.* **4,** 36–39.
20. Niedhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R., Kenyon, G. L. & Gerlt, J. A. (1991) *Biochemistry* **30,** 9264–9273.
21. Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. & Stevens, R. C. (1997) *Science* **276,** 1665–1669.
22. Ulrich, H. D., Mundorff, E., Santarsiero, B. D., Driggers, E. M., Stevens, R. C. & Schultz, P. G. (1997) *Nature (London)* **389,** 271–275.
23. Kaartinen, M., Solin, M. L. & Makela, O. (1991) *Eur. J. Immunol.* **21,** 2863–2869.
24. Xu, J., Deng, Q., Chen, J., Houk, K. N., Bartek, J., Hilvert, D. & Wilson, I. A. (1999) *Science* **286,** 2345–2348.
25. Jensen, R. A. (1976) *Annu. Rev. Microbiol.* **30,** 409–425.
26. Lewis, E. B. (1951) *Cold Spring Harbor Symp. Quant. Biol.* **16,** 159–163.