Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 3, pp 581–590, 1997

PII: S0968-0896(96)00271-4
0968-0896/97 $17.00 + 0.00

Expanding the 43C9 Class of Catalytic Antibodies Using a Chain-Shuffling Approach

Grover Paul Miller,[a] Bruce A. Posner[b] and Stephen J. Benkovic[a]*
[a]*Department of Chemistry, Pennsylvania State University, University Park, PA 16802, U.S.A.*
[b]*Pharmacology Department, University of Texas Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235, U.S.A.*

Abstract—We employed a chain-shuffling technique to determine if the light chain of the catalytic antibody, 43C9, provides the best partner for the 43C9 heavy chain. Previously, we reported construction and screening of a 43C9 HC CROSS library, where the 43C9 heavy-chain gene was crossed with a library of light-chain genes in a λ bacteriophage system. The library contained a high frequency of reconstituted antibodies recognizing the transition-state analogue. Here, we report the isolation and characterization of four of these clones. Recovered light-chain proteins share 92–96% sequence identity to the 43C9 light-chain protein. Somatic mutations of these light chains occur randomly at positions distant from the active site. Residues required for binding and catalysis were conserved. Mutations affected the topology of the binding site. Nevertheless, catalysis was not affected. Isolation of these light chains suggests the best partner for the 43C9 heavy chain is the original light chain. These clones attempt to broaden a class of 43C9-like antibodies, where the catalytic residues, His91 and Arg96, have been reproducibly selected. Similar catalytic properties between the 43C9-like antibodies suggests binding has been optimized, thus further maturation of the light chain would not lead to a better catalyst. To improve catalysis, other approaches must be considered.

Introduction

Enzymatic catalysis as outlined by J. B. S. Haldane[1] and elaborated by Linus Pauling results from the "complementarity of the enzyme's active site structure to the activated complex".[2] An immediate application of this concept has been the elicitation of catalytic antibodies. The strategy is to synthesize a stable mimic of the transition state of a desired reaction. Using this molecule as an immunogen yields antibodies that recognize and bind this transition-state analogue. These, in turn, are tested as putative catalysts for the chosen reaction. The approach has created the field of catalytic antibodies boasting catalysts for a wide array of different reactions.[3] To accomplish applications in biology, medicine, and biotechnology, we must increase the odds for their discovery, as well as their catalytic efficiency by pursuing a variety of methods for isolating and producing antibody catalysts.[4,5]

Confounding the discovery of catalytic antibodies is the immense repertoire of diverse antibodies, $>10^{10}$, generated in response to the immunogen.[6] They represent a formidable collection of antibodies to be screened as potential catalysts. The failure to isolate a catalytic antibody may derive from the size limitations imposed on the actual array of antibodies tested. The traditional approach employs hybridoma techniques that enable the researcher to screen several thousand monoclonals for recognition of a transition-state analogue. Of these, only a small population, typically 10–100, are assayed for catalysis, and represent a fraction of the total number of antibodies. Despite the limited number of candidates, the technique has successfully yielded a variety of catalysts. However, in no case, has the entire pool of antibodies been systematically explored.

The potential of this pool is underscored by the antibody 43C9, capable of accelerating the hydrolysis of a *p*-nitrophenyl-anilide 2.5×10^5 above the background rate, as well as cleaving a variety of aromatic ester.[7,8] The immunogen used to induce 43C9 was the *p*-nitrophenyl-phosphonamidate **4**, or NPN, in Figure 1. This transition-state analogue mimics the geometric and electronic characteristics of the transition state for *p*-nitrophenyl-anilide **1** as the molecule undergoes hydrolysis to yield *p*-nitroaniline **2** and the acid product **3**. Interestingly, the catalytic mechanism of 43C9 proved to be more complex than a simple complementation of the tetrahedral transition state by the antibody's active site. From modeling of the structure of 43C9 (confirmed by preliminary X-ray crystallographic analysis) and mutagenesis of specific amino acids, a light-chain residue, Arg96, was implicated in stabilizing the formation of the negative charge of the tetrahedral species. Another light-chain residue, His91, acts as a nucleophile to yield an acyl intermediate. Hydroxide ion presumably attacks the acyl intermediate to liberate a free antibody to continue the catalytic cycle.

To tap into the diverse repertoire of the immune system, another method, the chain-shuffling technique, exploits the architecture of the antibody, namely, the combination of heavy- and light-chains proteins to produce novel antibody combining sites.[9–10] After the immune response to a haptenic transition-state analogue, the

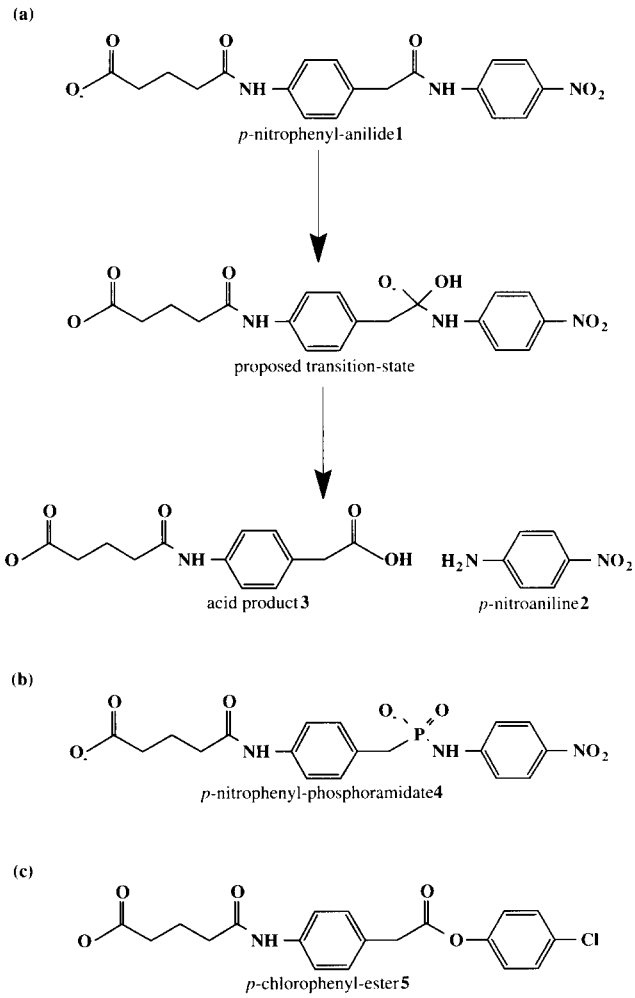

Figure 1. Molecules of interest. (a) Targeted reaction. To accomplish hydrolysis of *p*-nitrophenyl-anilide **1**, the carbonyl of the amide bond adopts a tetrahedral geometry with the concomitant creation of a negative charge. The passage of the molecule through this transition-state cleaves the amide to yield *p*-nitroaniline **2** and the acid product **3**. (b) Hapten. The *p*-nitrophenyl-phosphonamidate **4**, or NPN, mimics the geometric and electronic characteristics of the proposed transition-state of *p*-nitrophenyl-anilide **2**. (c) Substrate. To assay library antibodies for catalysis, *p*-chlorophenyl ester **5** was used as a substrate.

entire collection of antibody heavy- and light-chain genes can be recovered using the polymerase chain reaction. By in vitro techniques, the heavy- and light-chain genes are recombined in a λ bacteriophage-derived expression vector and the resulting combinatorial library is screened for clones with high affinity for the transition-state analogue. Selected clones are then assayed for catalysis. This combinatorial approach is not a new concept to studies involving antibodies. There are examples where the application of combinatorial antibody libraries has led to more diverse and higher-affinity antibodies for ligands,[11–14] and even a transition-state analogue.[15] The novelty in our approach is the search for a more effective catalytic antibody.

In a previous study, we crossed the 43C9 heavy-chain gene with a library of light-chain genes[16,17] to determine if the 43C9 light-chain protein possesses the most efficient catalytic contacts. We chose to retain the 43C9 heavy-chain gene for two reasons. First, holding one chain constant reduces the potential combinations from $>10^{10}$ to a number more effectively represented by a combinatorial library (typically about 10^6). Second, the 43C9 light-chain protein contains the catalytic machinery of the antibody catalyst. In this experiment, the 43C9 heavy-chain protein provides the foundation for the small substrate binding site, while the light-chain library offers a diverse array of sequences with potentially novel or more effective binding and catalytic residues than 43C9. The random combination of the 43C9 heavy-chain gene and a light-chain gene library derived from a mouse immunized with NPN hapten **4** yielded 2.6×10^7 members of the 43C9 HC CROSS library. Reconstituted antibodies were selected for recognition of NPN using a sandwich assay under stringent conditions ($K_d \leqslant 10$ nM). The population of 'binders' of the transition-state analogue proved large, about 24,500, versus the 44 hybridomas selected in the original study yielding 43C9.[9]

In this report, we describe the characterization of four randomly chosen 43C9 clones derived from the 43C9 HC CROSS library. The binding interactions were determined using NPN and the hydrolysis products, *p*-nitrophenol and the acid product. Catalytic activity was assayed using *p*-chlorophenyl-ester **5** as the substrate. The implications of these analyses, as well as the studies of other hydrolytic antibodies, are discussed in terms of the realistic potential of the immune system to yield effective hydrolytic catalysts.

Results

43C9 HC CROSS library yields 43C9-like clones

The 43C9 HC CROSS library proved to contain a high frequency of binders. From the Fab library of 2.6×10^7 members, one out of 1100 members recognized the transition-state analogue **4** with high affinity ($K_d \leqslant 10$ nM).[16] Four of these binders were randomly chosen for characterization to yield the 221 series. Figure 2 lists the amino acid sequences of the light-chain variable regions of clones from the screen, using the Kabat numbering system.[18] The 221 series shares 92–96% overall sequence identity to 43C9, representing essentially clones of 43C9. Such similarity permits the analysis of the recovered clones in terms of the well-characterized catalytic antibody 43C9. Notably, all possess the critical catalytic residues, His91 and Arg96. Differences in sequence reside in both complementarity determining regions (CDRs), L1, L2, and L3, and framework regions.

To visualize the locations of these amino acid substitutions, we took advantage of the structural model of the 43C9 Fv based on immunological sequence and structural data.[19] Figure 3 highlights the substituted positions of the clones superimposed on the 43C9 model. The wild-type sidechains at these positions are colored yellow against a blue ribbon diagram of the

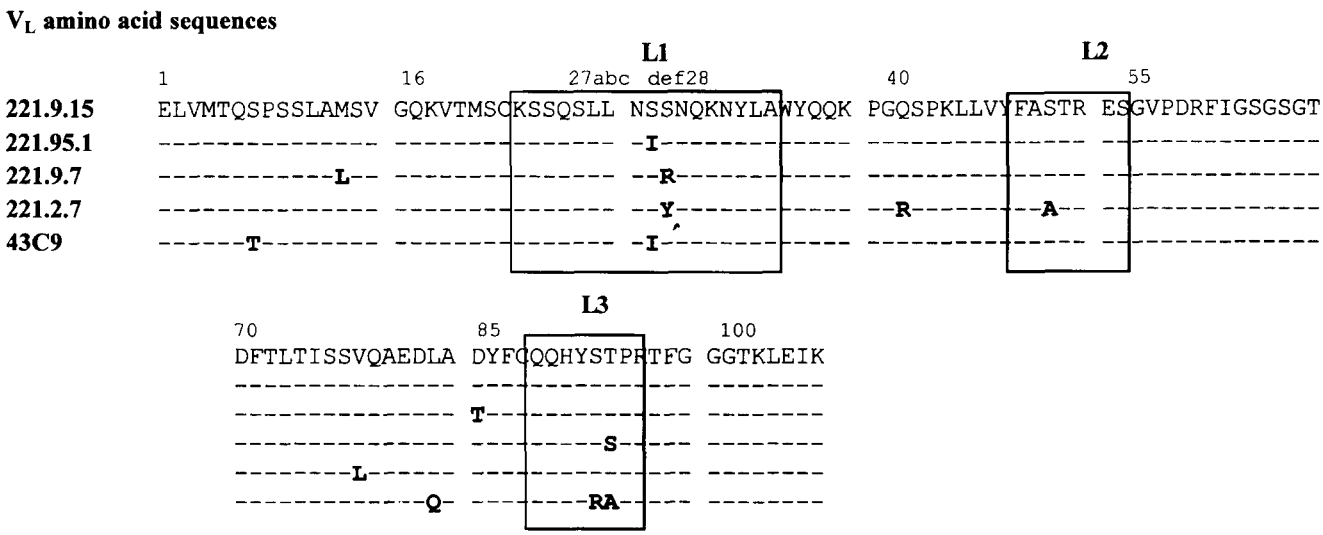

Figure 2. Comparison of the amino acid sequences of the variable light chains (V_L) of clones and 43C9. The 221 series reflects clones recovered from the library cross. Identical residues are indicated by a hyphen (-). Boldface type denotes the substitutions. CDRs (L1, L2, and L3) are represented by the boxed residues (43C9[7]).

light-chain protein. The respective heavy-chain protein is shown in red. The model illustrates that the differences in sequence occur at surface-exposed side-chains at a variety of distances from the active site occupied by the transition-state analogue **4**, NPN, in green in the figure. Although the 221 series light-chain proteins share high homology with 43C9, the amino acid substitutions are not conservative and occur in a number sufficient to prevent any quantitative predictions regarding binding and catalysis.

Analysis of variable regions

The origins of the light-chain genes were determined by comparison to known antibody sequences. The 221 series genes share two components: (1) a Vκ38 gene and (2) a Jκ1 element. No germline sequences have been reported for the Vκ8 class, despite an estimate of five to 16 sequences.[20] This combination of variable and J regions is common to a variety of antibody light-chain genes involved in recognizing typically small organic molecules. Figure 4 illustrates a comparison of the 221.9.15 variable light-chain amino acid sequence to sequences available from a BLAST search.[21] These light-chain proteins are recruited to participate in binding 2,4-dinitrophenyl,[22] DNA,[23] phosphorylcholine,[24] and glycophorin A type N blood group[25] antigens. All these genes differ from 221.9.15 mainly at position 96, where the Vκ gene combines with the Jκ region. The placement of arginine at this position is required for binding and catalysis by 43C9[26] and presumably by the clones.

From the nucleotide sequences in Figure 5, the 221.9.15 gene appears to represent a light-chain gene expressed before the 221.95.1, 221.9.7, and 221.2.7 genes. The 221.9.15 sequence bares high homology to VK139 differing at five positions resulting in three amino acid substitutions. Two of these substitutions exchange similar residues, a glutamic acid for an aspartic acid and a leucine for a valine, while the last change occurs at the aforementioned position 96. The significance of the similarity between these genes is the isolation of VK139 after the first days of birth for Balb/c mice,[22] suggesting this type of gene exists in the repertoire. The differences between 221.9.15 and the other 221 clones appears randomly, only occassionally introducing amino acid substitutions (bold type). These types of substitutions are typical of the hypermutation process late in the immune response.

Clones share similar characteristics

The nature of the library construct yielded antibody fragments (Fab), whose expression was poor and produced unstable proteins (data not shown). To overcome these drawbacks, the variable regions of the Fab genes were recovered by PCR and subcloned into the 43C9 scFv vector, pJS118. Since the clones share such high homology to 43C9, all binding and kinetic results are discussed as comparisons between the scFv form of 43C9 and the clones.

Thermodynamic dissociation constants

Catalysis by antibodies relies on the recognition of the transition-state analogue by the antibody. To assess this interaction for each antibody, the thermodynamic dissociation constant, K_d, for NPN hapten **4** was determined by following the quenching of intrinsic protein fluorescence (excitation λ = 290 nm; emission λ = 340 nm) upon ligand titration.[27] The results are listed in Table 1. All scFvs possessed high affinity for the NPN with K_D values below the detection limit of our technique (i.e., $K_d \leqslant 1$ nM). This finding is not surprising considering the high stringency of the library screen.

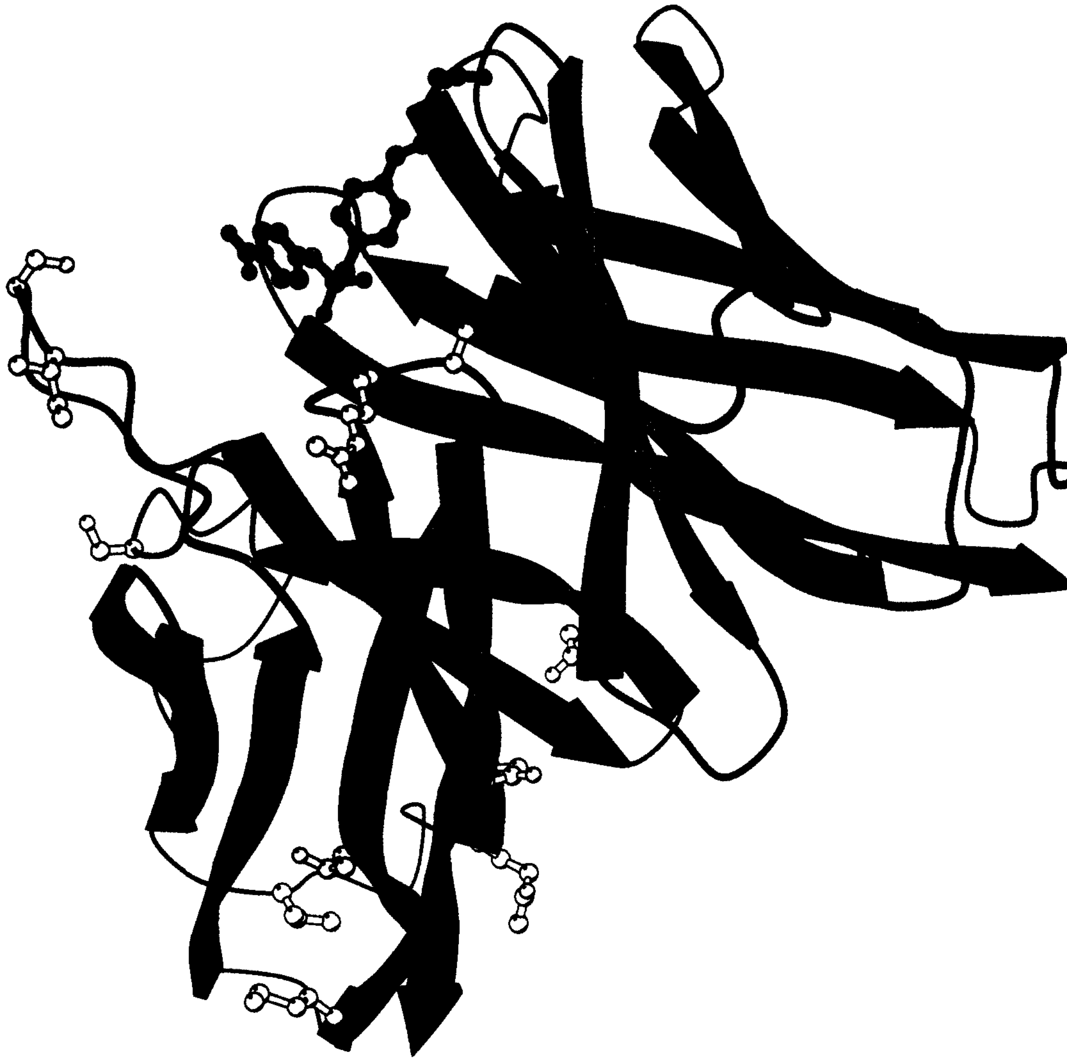

Figure 3. The structural model of the 43C9 Fv, where the blue and red ribbon diagrams are the light- and heavy-chain proteins, respectively. The NPN hapten **4**, shown in green, is bound in the active site. Sidechains colored yellow indicate positions where substitutions occur between the 221 series and 43C9.

V_L amino acid sequences

```
                                                                           L1
             antigen                          1                16        27abc   def28
221.9.15     p-nitrophenylphosphoramidate     ELVMTQSPSSLAMSV  GQKVTMSC KSSQSLL NSSNQKNYLAW YQQK
musigkab     2,4-dinitrophenyl                DI-------------  -------- ------- ----------- ----
mdigkvai     DNA                              DI-------------  -------- ------- ----------- ----
musigkaca    phosphorylcholine                DI-------------  -------- ------X ----------- ----
mmu01354     glycophorin A type N blood group ****-----------  -------- ------- ----------- ----

                      L2                                            L3
40                    55              70               85                    100
PGQSPKLLVY FASTR ESG VPDRFIGSGSGT DFTLTISSVQAEDLA DYFC QQHYSTPRT FG GGTKLEIK
---------- ----- --- ------------ --------------- ---- --------W -- --------
---------- ----- --- ------------ --------------- ---- --------Y -- --------
---------- ----- --- ------------ --------------- ---- --------P -- --------
---------- ----- --- ------------ --------------- V-Y- --------- -- --------
```

Figure 4. Comparison of the amino acid sequences of the 221.9.15 V_L with other antibodies. Identical residues are indicated by a hyphen (-). An asterisk (*) indicates the positions where no residues were reported and X is an unidentified residue. Boldface type denotes substitutions. The CDRs (L1, L2, and L3) are represented by the boxed residues. (musigkab,[22] mdigvai,[23] musigkaca,[24] mmu01354[25]).

Table 1. Thermodynamic dissociation constants of ligands for 221 series of antibodies

Antibody	Phosphoramidate 4 (nM)	Acid product 3 (μM)	*p*-Nitrophenol (μM)
43C9	<1	74.0 ± 5.7	0.154 ± 0.016
221.95.1	<1	377 ± 35	0.369 ± 0.030
221.9.15	<1	380 ± 61	0.397 ± 0.042
221.9.7	<1	350 ± 74	0.212 ± 0.023
221.2.7	<1	100 ± 25	0.466 ± 0.081

Table 2. Steady state parameters for *p*-chlorophenyl-ester **5** hydrolysis at pH 9.5 for the 221 series of antibodies

Antibody	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($mM^{-1}\,s^{-1}$)
43C9	0.960 ± 0.088	1.74 ± 0.27	0.550 ± 0.099
221.95.1	0.828 ± 0.067	1.80 ± 0.25	0.460 ± 0.074
221.9.15	0.938 ± 0.069	2.09 ± 0.24	0.449 ± 0.061
221.9.7	0.785 ± 0.043	1.29 ± 0.13	0.609 ± 0.070
221.2.7	0.746 ± 0.054	1.91 ± 0.23	0.390 ± 0.050

V_{L221} nucleotide sequences

```
                1               16              31              46              61              76
VK139 GACATTGTGATGACA CAGTCTCCATCCTCC CTGGCTATGTCAGTA GGACAGAAGGTCACT ATGAGCTGCAAGTCC AGTCAGAGCCTTTTA
221.9.15 --GC-C--------- --------------- --------------- --------------- ---------- ---------------
221.95.1 --GC-C--------- --------------- --------------- --------------- ---------- ---------------
221.9.7  --GC-C--------- --------------- ------C-------- --------------- ---------- ---------------
221.2.7  --GC-C--------- --------------- --------------- --------------- ---------- ---------------

                L1                                                                              L2
91              106             121             136             151             166
AATAGTAGCAATCAA AAGAACTATTTGGCC TGGTACCAGCAGAAA CCAGGACAGTCTCCT AAACTTCTGGTATAC TTTGCATCCACTAGG
--------------- --------------- --------------- --------------- --------------- ---------------
-------T------- --------------- --------------- --------------- --------------- ---------------
--------A------ --------------- --------A------ --------------- --------------- --------------A
------TA------- --------------- --------------G ------G-------- --------------- ---------------

181             196             211             226             241             256
GAATCTGGGGTCCCT GATCGCTTCATAGGC AGTGGATCTGGGACA GATTTCACTCTTACC ATCAGCAGTGTGCAG GCTGAAGACCTGGCA
--------------- --------------- --------------- --------------- --------------- ---------------
--------------- --------------- --------------- --------------- --------------- ---------------
--------------- --------------- --------------- --------------- --------------- ---------------
---G---------- --C------------ --------------- --------------- ---------T----- ---------------

                L3                                              Jκ1
271             286             301             316             331
GATTACTTCTGTCAG CAACATTATAGCACT CCGTGGACGTTCGGT GGAGGCACCAAGCTG GAAATCAAA
--------------- --------------- --TC------------ --------------- ---------
--------------- --------------- --TC------------ --------------- ---------
AC------------- --------------- --TC------------ --------------- ---------
--------------- -----------T-G- --TC------------ --------------- ---------
```

Figure 5. Nucleotide sequences of the 221 series light-chain genes compared to musigkab or VK139.[22] Identical bases are indicated by hyphens (-). Boldface type denotes substitutions that resulted in changes in the corresponding amino acid sequence. CDRs (L1, L2, and L3) are the boxed nucleotides. Although a definitive $V_κ8$ germline gene could not be determined, the $J_κ1$ element is underlined.

To understand the differences in binding between the clones, K_d values for products of *p*-nitrophenyl-ester hydrolysis were determined using the fluorescence quench technique. The acid product **3** and *p*-nitrophenol represent each half of the transition state analogue. The *p*-nitroaniline product was replaced by *p*-nitrophenol due to its better solubility under the assay conditions.

The binding between these molecules and the antibodies indicate the effects of the substitutions in the loop and framework regions on residues that contact regions of the substrate. The four amino acid substitutions of 221.95.1 resulted in a fivefold increase in the K_d for the acid product relative to 43C9. The K_d value was not affected by the additional substitution of Ile27eSer in 221.9.15, nor the changes in sequence for 221.2.7. Interestingly, the 221.9.7 K_d was similar to 43C9, indicating one or all the substitutions at positions 27f, 42, 52, and 94 contributed to the recovery of binding of the acid product. The *p*-nitrophenol K_d value for the 221 series was similar to 43C9. The changes in binding were modest for 221.95.1, 221.9.15, and 221.2.7. These clones experienced a two- to threefold decrease in binding for *p*-nitrophenol. The K_d value for 221.9.7 actually was similar to 43C9, perhaps due to the additional arginine at 27f in the L1 region.

Steady-state kinetics

To measure the catalytic activity of the clones, we chose the *p*-chlorophenyl-ester **5** substrate assaying its hydrolysis at pH 9.5. This substrate is one of many recognized by 43C9. The molecule is relatively easy to synthesize and provides a short assay time compared to the more suitable substrate, the *p*-nitrophenyl amide **1**. Further-

more, at pH 9.5 nucleophilic attack by His91 is rate limiting, thus any changes in the catalytic mechanism relative to 43C9 can be readily detected.

Overall, catalysis by the clones was similar to 43C9; however, there were interesting differences (Table 2). The k_{cat} values appeared to slightly decrease with an increasing number of substitutions away from the 43C9 sequence. The K_m values were equivalent within error for 43C9 and the various clones with the exception of 221.9.7, indicating that the binding interactions for the substrate were unchanged. The K_m for 221.9.7 was 1.4-fold less than 43C9, indicating a tighter binding of the substrate. This change may reflect the arginine at position 27f. Finally, with the exception of 221.9.7, which actually had a slightly higher catalytic efficiency than 43C9 resulting from the decreased K_m value, the catalytic efficiency for the clones, as measured by k_{cat}/K_m, slightly decreased with increasing substitutions.

Discussion

In this paper we investigated the potential of a light-chain library to provide a more catalytic light chain than the 43C9 light chain. Randomly combining the light-chain genes with the 43C9 heavy-chain gene produced a significant number of antibodies recognizing the transition-state analogue, phosphonamidate **4** (NPN), with high affinity. Four of these 'binders' (the 221 series) were isolated and characterized. Despite the use of a large pool of potentially diverse light-chain genes, all clones proved to share high sequence homology and display similar binding and catalysis as 43C9. The isolation of these clones suggests that the best partner for the 43C9 heavy chain is the original light chain. There are two possible reasons for this observation. First, the 43C9 heavy chain specifically recruited a 43C9-like light chain. Second, a 43C9-like light chain, in fact, possesses optimal binding contacts to the transition-state analogue. In either case, our results have implications regarding the potential of the immune system to yield effective hydrolytic catalysts.

Limited repertoire of hydrolytic antibodies

Antibodies recognizing small organic molecules, such as 2-phenyl-oxazolone,[28] have been shown to involve a restricted population of genes. The same trend presumably applies to studies of catalytic antibodies utilizing small organic molecules as transition-state mimics, such as in this study. The additional requirement of catalysis further limits the pool of antibodies. Nevertheless, a variety of unique antibodies have been isolated that catalyze hydrolysis of ester substrates. These antibodies conceivably represent sequence-based classes of antibody catalysts. The isolation of 43C9-like light chain protein in this study that complement the 43C9 heavy-chain protein attempts to broaden one of these classes.

The origins of 43C9 indicate the possibility of a 43C9-like class of antibodies. This antibody was isolated late in the immune response of Balb/c mice to the transition-state analogue, NPN, suggesting 43C9 is a matured antibody. The high affinity of this antibody for NPN ($K_D \leqslant 10$ nM) agrees with this conclusion. To achieve this affinity, the evolution of a binder would probably involve a very specific partner. As maturation progresses, the immune system systematically creates and screens the best pair of chains. The light-chain library used to construct the 43C9 HC CROSS library derives also from Balb/c mice at a matured point in antibody production. This pool of genes potentially contains the products of the evolution of the 43C9 light chain.

Although the original 43C9 light chain was not found in this study, its presence in the library is statistically possible. Estimates of antibody-chain pairings involve the following equation[29,30]

$$P = 1 - e^{-q}$$

where $q = N \times P_H \times P_L$.

An assumption of this equation is that the probability (P) of a unique heavy- and light-chain pair occurring in a combinatorial library is determined by a Poisson distribution.[29] The size of the library is represented by N, and the parameters P_H and P_L are the probabilities of random occurrence of the heavy- and light-chain gene fragments, respectively. In this experiment, the size of the library (N) is 2.6×10^7. Since the heavy-chain gene is constant, the parameter P_H equals one, and the variable P_L was experimentally determined to be 9.0×10^{-4} from a previous study.[16] If we apply these values to the equation, the probability of finding the original 43C9 heavy- and light-chain combination is >99%.

Although a systematic sequencing of the library was not performed in this study, the high sequence identity of the 221 series to 43C9 suggests the respective light-chain gene may be present in the pool of light-chain genes. Furthermore, a hybridization screen of the light-chain library revealed a significant number of light-chains may be '43C9-like'.[17] The study employed an oligonucleotide complementary to the CDR3 region of the 43C9 light-chain. Under high stringency, one in 200 sequences were indicated as possessing this region of the light-chain known to be critical in binding and catalysis. In following, it is plausible to isolate 43C9-like light chains.

The isolation of four 43C9-like light chains indicates that the 43C9 heavy chain may be selecting out this type of light chain. Further evidence for this trait is found in a complementary shuffle.[16,17] In that study a combinatorial library was made by crossing the 43C9 light-chain gene with a heavy-chain gene library derived from a Balb/c mouse immunized with the transition-state analogue, NPN. The resulting 43C9 LC CROSS library of reconstituted antibodies was screened for recognition

of NPN, where $K_D \leqslant 10$ nM. Three 'binders' from this library proved to contain heavy-chain sequences divergent from each other, as well as from the 43C9 heavy-chain gene. These results are consistent with a higher stringency of the light chain relative to the heavy chain, a trend also observed in two classes of esterolytic monoclonal antibodies.[31–33]

Other groups have reported similar findings from studies of esterolytic monoclonal antibodies. One study included six antibodies capable of hydrolyzing a *p*-nitrophenyl-ester substrate, as well as a carbonate substrate by some antibodies.[31] These catalytic antibodies fell into three sequence-based classes. In one class three antibodies shared about 91–93% and 99% sequence homology in the heavy- and light-chain proteins, respectively. The other two classes shared a unique light-chain gene, but contained different heavy-chain genes. Another group found five out of six antibodies hydrolyzing an ester substrate showed 74–84% and 89–95% sequence in the heavy- and light-chain proteins, respectively.[32] A third study yielded esterolytic antibodies with 84–91% sequence homology between the heavy-chain proteins and at least 94% sequence homology between the light-chain proteins.[33] In contrast to the these reports, we obtained a similar result employing an in vitro chain-shuffling approach rather than the hybridoma technique yielding monoclonal antibodies. Our 221 series of 43C9-like antibodies represent another sequence-based class of catalysts.

Maturing the immune repertoire

As the immune response to an antigen progresses, antibodies display an increasing number of mutations presumably to maximize recognition for a particular antigen. This feat is brought about through the diversification of the repertoire of antibody genes. The isolated light-chain genes of this study are products of the maturation process. A potential progenitor for the 221 series genes is the VK139 light-chain gene, which was isolated after the first days of birth for Balb/c mice. This suggests that the 43C9-like type of gene may already exist in the repertoire. Comparing the 221.9.15 nucleotide sequence to VK139 in Figure 4 reveals five differences that result in three amino acid substitutions. The only significant change in sequence occurs at the junction of the Vκ8 and Jκ 1 elements, where 221.9.15 has an arginine and VK139 has a tryptophan. The VK139 light-chain gene is from a clonotype recognizing 2,4-dinitrophenyl antigen. The clonotype predominates in the first week of neonatal Balb/c mice,[22] thus the immune repertoire contains a 221.9.15 gene or a germline sequence resembling it.

Analysis of 221.95.1, 221.9.7, and 221.2.7 suggest these light-chain genes derive from a matured immune response. From the nucleotide sequences of the 221 series in Figure 4, the 221.95.1, 221.9.7, and 221.2.7 genes contain a variety of substitutions compared to 221.9.15. The changes in sequence appear in both CDR and framework regions. Roughly half of these result in a change in the amino acid sequence (bold type). Notably, there are no substitutions in the active site, as indicated by the 43C9 Fv model (Fig. 3). In other words, the residues contacting the transition-state analogue have been optimized. Possibly, these amino acids are involved in subtle substructural changes in the binding surface, which result in greater molecular flexibility of the antibody molecule. The variations in the K_D values for the end-products indicate that the topology of the binding site is affected by these distant mutations. A recent study reported a similar finding.[34] Random substitution in the antibody gene is characteristic of the hypermutation mechanism acting on germline sequences to diversify the primary repertoire.[35–37] Products of this process are isolated in the second week of the immune response to small organic haptens.[38–40] Consequently, the 221.95.1, 221.9.7, and 221.2.7 genes arise from the maturation of 221.9.15 or a similar sequence.

A common difference between the catalytic light-chain amino acid sequences and other similar sequences (Fig. 4) occurs at position 96, where the 221 series and 43C9 have an arginine. Shown in Figure 3, a model of the 43C9 Fv suggests the residue interacts electrostatically with the negative charge on the phosphoryl portion of the hapten. The placement of an arginine at the VκJκ junction is observed in 48 of 48 phenylarsonate-specific antibodies.[41] Furthermore, another phosphonate hapten used in a separate study resulted in an esterolytic antibody possessing an arginine in the 96 position of the light-chain protein.[34] The remainder of reported esterolytic antibodies lack an arginine at this position, although one study did report a lysine involved in charge complementation of the hapten from position 97 of the heavy-chain protein.[33] The catalytic efficiency of 43C9 may reflect the strategic placement of an arginine within the active site. Mutagenesis of the arginine in 43C9 to glutamine did reduce binding to NPN, but did not affect binding to the end-products of hydrolysis,[26] thus supporting the predicted role in charge complementation. The loss of the arginine also eliminated the catalysis by the antibody. Because all catalytic light-chain sequences contain the arginine, the phosphoryl portion of the transition-state analogue presumably plays a role in selecting for the catalytically-important arginine.

Another residue crucial for catalysis, His91, is found in all isolated light chains and 43C9. The presence of His91 results from two possible sources. The histidine may be present in a germline sequence. This residue is simply not mutated during maturation of the gene, thus the isolation of light-chain sequences containing His91 is fortuitous. Otherwise, His91 of the 43C9 light chain plays a role in binding, providing a selective factor for antibodies possessing this residue. Although the role of His91 in binding has not been studied by site-directed mutagenesis, a model of the 43C9 Fv places the imidazole sidechain in proximity to the phosphoryl portion of NPN.[19] A histidine adjacent to a phosphonate transition-state analogue has also been observed

with three different esterolytic antibodies.[33,34,42] In contrast to 43C9, the histidine of these studies is at position 35 of the heavy chain. The residue at position 91 of the respective light chains is tyrosine or glycine. This difference in the placement of the histidine may explain why 43C9 is a more potent catalyst, capable of hydrolyzing amides and esters.

The origin of the 43C9 heavy-chain gene is more difficult to determine, since the heavy-chain gene derives from three genetic elements: V_H, D, and J_H. Moreover, the D element may be partially or completely deleted, as well as inverted. The combination of these elements can result in low overall sequence homology between heavy-chain genes derived from the same elements. This observation reflects high variability in CDR3, the point where all genetic elements combine. Nevertheless, a monoclonal esterolytic antibody, CNJ123, was isolated that bore sequence homology to 43C9.[31] The heavy-chain protein of CNJ123 proved to share 74% overall sequence homology with the 43C9 heavy-chain protein. A major difference between the sequences resides in the CDR3 region, where the D element of CNJ123 has been essentially deleted. The source of the homology between the full length genes is the respective V_H and J_H elements. The combination of these types of elements has been found in heavy-chain genes of antibodies binding the small organic molecule, 2-phenyl-oxazolone.[28] The similarity of the resulting heavy-chain proteins suggest some structural contacts are common between the respective heavy- and light-chain proteins.

Evolution of a catalyst

A cornerstone of the catalytic antibody field is that the immune system provides adequate material to construct a catalyst. The principle assumes that a sufficient variety of antibody scaffolds exist to provide a foundation for a catalyst and that the residues within the binding site can be optimized for binding a transition-state analogue to yield catalysts. Assuming the 43C9-like light-chain proteins combined with the 43C9 heavy-chain protein represent a scaffold for a catalytic antibody, the sequences of these light-chain proteins provide insight into the evolution of this group of antibodies as catalysts. These antibodies displayed similar catalytic properties. Although the 221 series of light-chain proteins represent a small portion of the potential catalysts in the HC CROSS Library (4 out of about 24,500), these results raise the question as to whether the immune system can provide a more potent catalyst than 43C9.

The design of the NPN hapten may have contributed to the limit of catalysis by 43C9. The transition-state analogue utilizes the tetrahedral, negatively-charged phosphoryl center to mimic the purported transition-state of an amide or ester undergoing hydrolysis. The success of a phosphonate structure is well-documented in studies of protease inhibitors.[43,44] The high affinity of these inhibitors typically arises from strong hydrogen bonds. These interactions may form due to the larger size of the phosphonate relative to the proposed transition-state. The possible ceiling limiting catalysis by antibodies may reflect the appropriateness of phosphonates as true transition-state mimics.

There are examples of other protease inhibitors that may prove to be interesting haptens for studies of catalytic antibodies. Sulfonates and hydrated ketones can mimic the geometric transition-state of hydrolyzing esters and amides, although these molecules lack a negative charge. The use of these types of molecules as haptens would potentially yield smaller binding pockets that more effectively desolvate the substrate, a tactic employed by a variety of enzymes. In this case, one of the advantages of the chain-shuffling technique becomes apparent. It is possible to create antibody-chain libraries derived from different haptens and cross them to yield antibodies combining the traits of the individual libraries.

The light-chain proteins of this study are matured products of the immune system, as evidenced by the high affinity for the transition-state analogue. In a landmark study, one group was able to correlate the maturation of an antibody to binding a transition-state analogue to an increase in catalysis.[34] In following, the matured light-chains of this study have been optimized for binding NPN, since there was no significant effect on catalysis. The somatic mutations amongst the light chains only resulted in distal mutations potentially affecting the topology of the antibody. At least with the 43C9 construct, we may have reached the limit of catalysis from nature's process of random mutagenesis.

Future strategies

Although we did not isolate a more effective antibody catalyst than 43C9, we gained insight into the utility of the chain-shuffling approach, as well as the ability of the immunological repertoire to provide catalysts. The 43C9 HC CROSS library contained a formidable number (approximately 24,500) of antibodies recognizing the transition-state analogue. Nevertheless, there exists a restricted number of heavy- and light-chain protein combinations yielding hydrolytic antibodies. When antibody genes are optimized for binding during maturation, catalysis presumably reaches an endpoint. Although a thorough examination of all potential catalysts was not undertaken in this study, the isolation of 43C9-like antibodies displaying similar characteristics serves to broaden a class of antibodies and may reflect the limit of catalysis by a 43C9 construct. To go beyond the 'natural' limit on catalysis, other approaches should also be considered, such as site-directed mutagenesis, protein engineering, or chain-shuffling of libraries derived from different haptens.

Experimental

General

Standard DNA and bacterial procedures were followed as described.[45] Propagation of plasmid DNA was carried out using *Escherichia coli* DH5α cells. The expression strain of *E. coli*, BL21 (genotype *hsdS*, *gal*) and the T3 expression plasmid, pTG119, were generous gifts from Robert LaPolla (R. W. Johnson Pharmaceutical Research Institute). Restriction and DNA-modifying enzymes were purchased from New England Biolabs, Boehringer-Mannheim, or Promega. DNA sequences were determined using the Sequenase kit (U.S. Biochemical). Oligonucleotides for mutagenesis and sequencing purposes were obtained from Integrated DNA Technologies, Inc. All materials for the polymerase chain reaction were obtained through Promega. The synthesis of *p*-nitrophenyl-phosphoramidate **4**, *p*-chlorophenyl-ester **5**, and acid product **3** were performed as previously described.[7,46]

Fab clones converted to scFv form

The variable regions of Fab light-chain genes were rescued by PCR, restriction digested, and then subcloned into a vector containing the remainder of the scFv gene. The target vector was pJS118,[19] which contains the 43C9 scFv gene. A *Sac*I restriction site lies at the 5′ end of the light chain variable gene, while a *Hin*dIII restriction site lies at the 3′ end. To replace the wild-type 43C9 light-chain gene, clones required these restriction sites. Although the *Sac*I site is present in the Fab construct, a *Hin*dIII site had to be introduced. This was accomplished by using a mutagenic primer and PCR mutagenesis. The relevant primers were:

For primer:
$^{5'}$AAACTAGTCGCCAAGGAGACAGT$^{3'}$

Rev primer:
$^{5'}$TTTTATTTCAAGCTTGGTGCCTCC$^{3'}$

*Hin*dIII

After the genes were amplified, they were digested with *Sac*I and *Hin*dIII and subsequently gel purified. The insert was then ligated into the gel-purified vector, previously isolated from the 43C9 light-chain gene. To ensure substitutions were not introduced during the PCR step, the resultant constructs were sequenced and compared to the original Fab light-chain variable genes.

Isolation of scFv protein

A typical isolation of a scFv involved the modification of a published procedure.[19] The expression system employed two plasmids. One contained the scFv gene under the control of the *lac* promoter. The other plasmid, pTG119, contained a *lac* i_q gene and a T3 RNA polymerase gene downstream of a *lacUVR* promoter. Both plasmids were used in the transformation of BL21 (DE3) *Escherichia coli* cells. Cells were grown in the presence of 10 μg mL^{-1} tetracycline and 25 μg mL^{-1} kanamycin, then subcultured into 4 L of media. After an overnight induction with 200 μM IPTG for 16–18 h, the cells were harvested and the supernatant concentrated. The antibody was purified from the supernatant and dialyzed into a storage buffer (100 mM HEPES, 50 mM NaCl, pH 7.5) at 4 °C. Typical yields ranged from 25 to 900 μg of scFv from a 4 L preparation.

Fluorescence titrations of scFvs with various ligands

To obtain thermodynamic dissociation constants (K_d) for ligands, the quenching of intrinsic protein fluorescence was measured as a function of ligand concentration using an SLM Aminco 8000 spectrofluorimeter (SLM Instruments, Inc.). Typically, the antibody was added to a filtered, degassed buffer solution of 100 mM HEPES, 50 mM NaCl, pH 7.5 in a fluorescence cuvette. The excitation wavelength was 290 nm and the emission wavelength was 340 nm. The addition of ligand resulted in a decrease in fluorescence. After correction for inner filter effects, the data were fit to a quadratic equation.[27] Acceptable K_d determinations were made using concentrations below the respective K_d value. To quantitate the antibody, the active site was titrated with hapten **4**, using the quenching of fluorescence as above. If protein concentration exceeds by 10-fold the K_d for the ligand, a monotonic decrease in fluorescence will be observed. A fit of these data yields the concentation of active antibody.

Steady-state kinetic parameters

The standard kinetic assay involved monitoring the hydrolysis of the *p*-chlorophenyl-ester substrate **5** in Scheme 1 as the increase in absorbance at 282 nm (isobestic point of *p*-chlorophenol; $\varepsilon = 1392$ M^{-1}) using a Cary UV-vis spectrometer (Varian Instruments, Inc.). The reaction buffer contained 0.8 μM scFv and 2.5% (vol/vol) dimethylformamide in filtered, degassed buffer, 100 mM CHES, 25 mM NaCl at pH 9.5. Each assay was initiated upon addition of substrate at 25 °C. The concentration of substrate was varied from 200 to 2000 μM *p*-chloro-phenyl-ester, the limit of solubility. The initial velocities were measured and corrected for background rates. To obtain K_m and k_{cat} values, the data were fitted to a Michaelis–Menton kinetic scheme.

Acknowledgments

We gratefully acknowledge Dr Craig E. Cameron for helpful comments and valuable discussions. G. P. M. is a recipient of a Homer F. Braddock College of Science Memorial Scholarship.

References

1. Haldane, J. B. S. In *Enzymes*; Green: London, 1930; pp 179–185.

2. Pauling, L. *Chem. Eng. News* **1946**, *24*, 1375.

3. Lerner, R. A.; Benkovic, S. J; Schultz, P. G. *Science* **1991**, *252*, 659.

4. Tawfik, D. S.; Eshhar, Z.; Green, B. S. *Mol. Biotech.* **1994**, *1*, 87.

5. Benkovic, S. J. *Annu. Rev. Biochem.* **1992**, *61*, 29.

6. Winter, G.; Milstein, C. *Nature (London)* **1991**, *349*, 293.

7. Janda, K. D.; Schloeder, D.; Benkovic, S. J.; Lerner, R. A. *Science* **1988**, *241*, 1188.

8. Stewart, J. D.; Krebs, J. F.; Siuzdak, G.; Berdis, A. J.; Smithrud, D. B.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7404.

9. Huse, W. D.; Sastry, L.; Iverson, S. A.; Kang, A. S.; Alting-Mees, M.; Burton, D. R.; Benkovic, S. J.; Lerner, R. A. *Science* **1989**, *246*, 1275.

10. Posner, B.; Lee, I.; Itoh, T.; Pyati, J.; Graff, R.; Thorton, G. B.; LaPolla, R.; Benkovic, S. J. *Gene* **1993**, *128*, 111.

11. Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. *Nature (London)* **1991**, *352*, 624.

12. Kang, A. S.; Barbas, C. F.; Janda, K. D.; Benkovic, S. J.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4363.

13. Collet, T. A.; Roben, P.; O'Kennedy, R.; Barbas, C. F.; Burton, D. R.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10026.

14. Marks, J. D.; Griffiths, A. D.; Malmqvist, M.; Clackson, T. P.; Bye, J. M.; Winter, G. *Bio/Technology* **1992**, *10*, 779.

15. McCafferty, J.; Fitzgerald, K. J.; Earnshaw, J.; Chiswell, D. J.; Link, J.; Smith, R.; Kenten, J. *Applied Biochem. Biotech.* **1994**, *47*, 157.

16. Posner, B. P. In *Approaches to Engineering and Characterizing Novel Biological Catalysts*. Ph.D. dissertation, Pennsylvania State University, 1994; pp 121–300.

17. Posner, B.; Smiley, J.; Lee, I.; Benkovic, S. *TIBS* **1994**, *19*, 145.

18. Kabat, E. A.; Wu, T. T.; Reid-Miller, M.; Perry, H.; Gottsman, K. S. *Sequences of Proteins of Immunological Interest*. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health: Bethesda, 1987.

19. Roberts, V. A.; Stewart, J.; Benkovic, S. J.; Getzoff, E. D. *J. Mol. Biol.* **1994**, *235*, 1098.

20. Strohal, R.; Helmberg, A.; Kroemer, G.; Kofler, R. *Immunogenetics* **1989**, *30*, 475.

21. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.

22. Riley, S. C.; Connors, S. J.; Klinman, N. R.; Ogata, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2589.

23. Tillman, D. M.; Jou, N.-T.; Hill, R. J.; Marion, T. N. *J. Exp. Med.* **1993**, *176*, 761.

24. Pennel, C. A.; Maynard, E.; Arnold, L. W.; Haughton, G.; Clarke, S. H. *J. Immunol.* **1990**, *145*, 1592.

25. Czerwinski, M.; Blackall, D. P.; Abrams, W. R.; Rubocki, R. J.; Spitalnik, S. L. *Mol. Immunol.* **1994**, *31*, 279.

26. Stewart, J. D.; Roberts, V. A.; Thomas, N. R.; Getzoff, E. D.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 1994.

27. Taira, K.; Benkovic, S. J. *J. Med. Chem.* **1988**, *31*, 129.

28. Kaartinen, M.; Griffith, G. M.; Markham, A. F.; Milstein, C. *Nature (London)* **1983**, *304*, 320.

29. Perelson, A. S.; Oster, G. F. *J. Theor. Biol.* **1979**, *81*, 645.

30. Perelson, A. S. *Immunol. Rev.* **1989**, *110*, 5.

31. Zemel, R.; Schindler, D. G.; Tawfik, D. S.; Eshhar, Z.; Green, B. S. *Mol. Immunol.* **1994**, *31*, 127.

32. Miyashita, H.; Hara, T.; Tanimura, R.; Tanaka, F.; Kikuchi, M.; Fujii, I. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6045.

33. Guo, J.; Huang, W.; Zhou, G. W.; Fletterick, R. J.; Scanlan, T. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1694.

34. Patten, P. A.; Gray, N. S.; Yang, P. L.; Marks, C. B.; Wedemayer, G. J.; Boniface, J. J.; Stevens, R. C.; Schultz, P. G. *Science* **1996**, *271*, 1086.

35. Allen, D.; Cumano, A.; Dildrop, R.; Kocks, C.; Rajewsky, K.; Rajewsky, N.; Roes, J.; Siekevitz, M. *Immunol. Rev.* **1987**, *96*, 5.

36. Berek, C.; Milstein, C. *Immunol. Rev.* **1987**, *96*, 23.

37. Sharon, J.; Gefter, M. L.; Wysocki, L. J.; Margolies, M. N. *J. Immunol.* **1989**, *142*, 596.

38. Weiss, U.; Zoebelein, R.; Rajewsky, K. *Eur. J. Immunol.* **1992**, *22*, 511.

39. Griffiths, G. M.; Berek, C.; Kaartinen, M.; Milstein, C. *Nature (London)* **1984**, *312*, 271.

40. Weiss, U.; Rajewsky, K. *J. Exp. Med.* **1990**, *172*, 1681.

41. Jeske, D. J.; Jarvis, J.; Milstein, C.; Capra, J. D. *J. Immunol.* **1984**, *133*, 1090.

42. Golinelli-Pimpaneau, B.; Gigant, B.; Bizebard, T.; Navaza, J.; Saludjian, P.; Zemel, R.; Tawfik, D. S.; Eshhar, Z.; Green, B. S.; Knossow, M. *Structure* **1994**, *2*, 175.

43. Phillips, M. A.; Kaplan, A. P.; Rutter, W. J.; Bartlett, P. A. *Biochemistry* **1992**, *31*, 959.

44. Fraser, M. E.; Strynadka, N. C. J.; Bartlett, P. A.; Hanson, J. E.; James, N. G. *Biochemistry* **1992**, *31*, 5201.

45. Sambrook, J.; Fritsch, E. F.; Maniatus, T. *Molecular Cloning. A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, 1989.

46. Gibbs, R. A.; Benkovic, P. A.; Janda, K. D.; Lerner, R. A.; Benkovic, S. J. *J. Amer. Chem. Soc.* **1992**, *114*, 3528.

(*Received in U.S.A.* 26 *June* 1996; *accepted* 5 *November* 1996)