# Antibody evolution constrains conformational heterogeneity by tailoring protein dynamics

Jörg Zimmermann*, Erin L. Oakman*, Ian F. Thorpe†, Xinghua Shi‡, Paul Abbyad‡, Charles L. Brooks III†, Steven G. Boxer‡, and Floyd E. Romesberg*§

Departments of *Chemistry and †Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037; and ‡Department of Chemistry, Stanford University, Stanford, CA 94305

Edited by Peter G. Wolynes, University of California at San Diego, La Jolla, CA, and approved July 21, 2006 (received for review April 21, 2006)

**The evolution of proteins with novel function is thought to start from precursor proteins that are conformationally heterogeneous. The corresponding genes may be duplicated and then mutated to select and optimize a specific conformation. However, testing this idea has been difficult because of the challenge of quantifying protein flexibility and conformational heterogeneity as a function of evolution. Here, we report the characterization of protein heterogeneity and dynamics as a function of evolution for the antifluorescein antibody 4-4-20. Using nonlinear laser spectroscopy, surface plasmon resonance, and molecular dynamics simulations, we demonstrate that evolution localized the Ab-combining site from a heterogeneous ensemble of conformations to a single conformation by introducing mutations that act cooperatively and over significant distances to rigidify the protein. This study demonstrates how protein dynamics may be tailored by evolution and has important implications for our understanding of how novel protein functions are evolved.**

flexibility | nonlinear spectroscopy | fluorscein | molecular recognition

Modern theories of protein evolution suggest that the most efficient pathway to evolve proteins with new function starts with precursor proteins that are flexible or conformationally heterogeneous (1–3). The precursor proteins are able to adopt multiple conformations, in addition to the one that is optimal for their primary function. If a rare conformation is suitable for a different and beneficial activity, there is an immediate selective advantage to duplication of the corresponding gene, which may then acquire mutations that stabilize and optimize the rare conformation.

The paradigm of these theories is the immune system, wherein mature Abs specific for virtually any foreign molecule are rapidly evolved from a limited set of precursor (or "germ-line") Abs. To accomplish this feat of molecular recognition, it has been suggested that the repertoire of germ-line Abs may have been selected to be flexible and/or conformationally heterogeneous to ensure recognition of the broadest range of target molecules (4–9). Although these flexible, polyspecific germ-line Abs are also expected to recognize self molecules (10), they are not present at concentrations sufficient to cause autoimmunity (11). Abs specific for a foreign molecule may then be evolved when a rapid change in concentration or presentation of the foreign molecule triggers a mutagenic proliferation of the germ-line Ab (12, 13). During this process, known as somatic evolution, mutations may be selected that simultaneously increase affinity and selectivity if they act, at least in part, to restrict the Ab to a conformation that is appropriate for recognition of the foreign molecule (8, 10, 11, 14–21). The resulting Abs are specific for their foreign targets and thus may be produced at increased levels without risk of self-recognition and autoimmunity. Thus, conformational restriction might underlie the evolution of mature Abs from germ-line Abs. Although this mechanism of Ab evolution has been widely cited, there is virtually no direct experimental evidence that flexibility or conformational heterogeneity of an Ab, or any other protein, may be optimized during evolution.

To test the hypothesis that evolution restricts Ab flexibility and/or conformational heterogeneity, the specific mutations introduced during evolution must be determined. Germ-line Abs are assembled from a set of known genomic fragments, which may be determined by comparing the 5′ UTR of candidate genomic fragments with that of the rearranged genes (17). Mutations identified by comparing these sequences are typically found throughout the Ab-combining site, which is formed from the six loops or complementarity-determining regions (CDRs) that connect the strands of the $\beta$-sheet framework (Fig. 1). Three CDRs are provided by the variable region of a light-chain polypeptide ($V_L$ CDR1-3) and three by the variable region of a heavy-chain polypeptide ($V_H$ CDR1-3). Particularly elegant studies by Wedemayer *et al.* (8) and Patten *et al.* (17) showed that somatic mutations throughout the Ab-combining site may preorganize the CDRs for binding. In addition, thermodynamic studies have shown that germ-line Abs may bind their targets with a more negative entropy, relative to mature Abs (22, 23). Although these results are consistent with the model that affinity maturation transforms flexible receptors into more rigid receptors, the studies did not actually measure flexibility or conformational heterogeneity.

To characterize the relationship between evolution, flexibility, and conformational heterogeneity, a quantitative measure of flexibility and heterogeneity is required. Generally, conformational heterogeneity may be described according to Frauenfelder's model of a hierarchical energy landscape wherein proteins exist in different conformations, with each conformation consisting of a large number of conformational substates (CSs) (24). Protein flexibility results from fluctuations between CSs that occur on the ps to ns time scale, and conformational heterogeneity results from transitions between different conformations that occur on longer time scales (25, 26). One approach to experimentally characterizing protein flexibility is based on measuring how a protein relaxes after displacement from equilibrium by a photoinduced change in the charge distribution of a bound chromophore (27–30). The induced motions are manifest as discrete peaks in the Ab spectral density, $\rho_{Ab}(\omega)$, which is the frequency domain representation of the ensemble-averaged time-correlation function of the electronic transition energy gap, $M(t)$ (31). $\rho_{Ab}(\omega)$ thus describes the amplitude of protein motions as a function of their frequency and thus may be used to characterize flexibility. Flexible proteins exhibit low-frequency amplitude, whereas more rigid proteins exhibit increased high-frequency amplitude (29). $\rho_{Ab}(\omega)$ may be deter-

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: Fl, fluorescein; CDR, complementarity-determining region; 3PEPS, three-pulse photon echo shift; DSS, dynamic Stokes shift; MD, molecular dynamics.

§To whom correspondence should be addressed. E-mail: floyd@scripps.edu.





www.pnas.org/cgi/doi/10.1073/pnas.0603282103

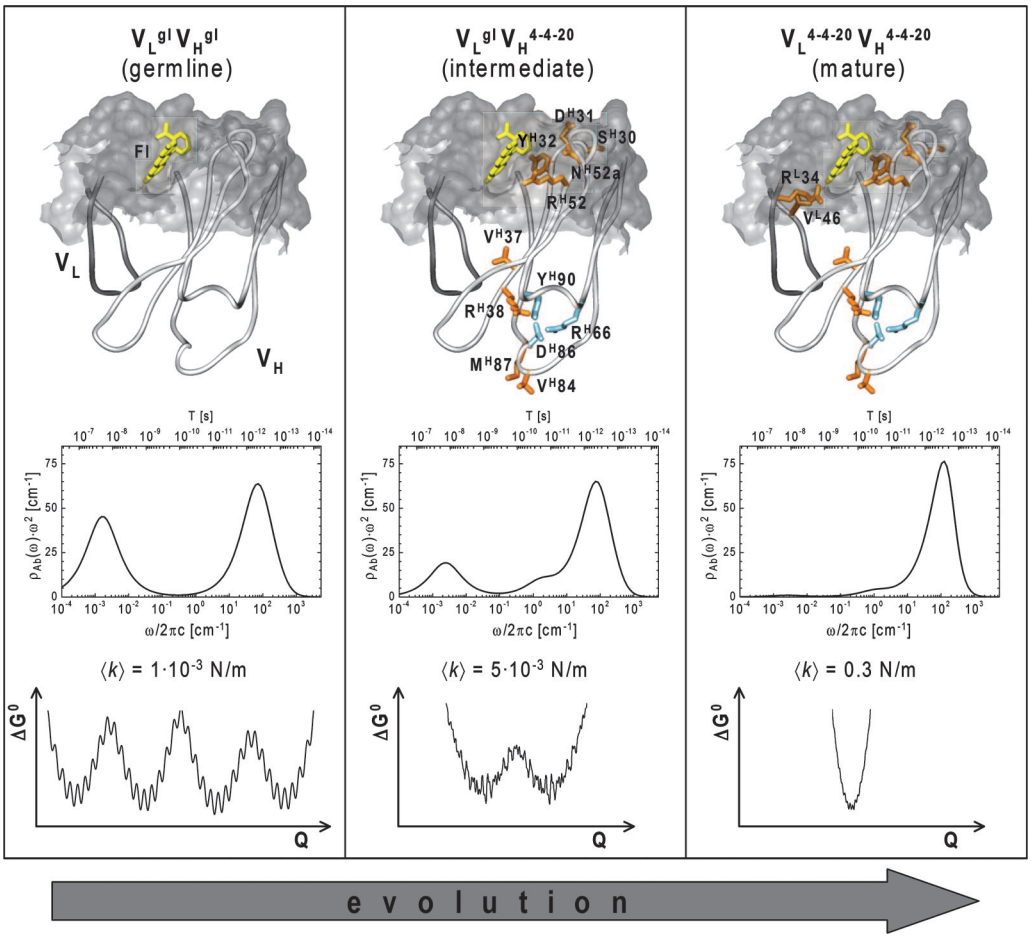


**Fig. 1.** Evolution of protein structure and dynamics of Ab 4-4-20. (*Top*) Ab variable regions showing the mutations introduced during Ab evolution (orange). Also shown are the residues that form the hydrogen-bond network with $R^H38$ in the mature Ab (light blue). For clarity, only part of the light chain is shown. (*Middle*) Spectral densities for the three proteins from 3PEPS and DSS data. The corresponding oscillation period, $T = 2\pi/\omega$, is included for comparison. (*Bottom*) Schematic representation of the Ab energy landscape at various stages of evolution. $\langle k \rangle$ is the average force constant of the combining site determined from 3PEPS and DSS experiments.

mined from three-pulse photon echo peak shift (3PEPS) experiments (27) or dynamic Stokes-shift (DSS) experiments (32, 33). Importantly, 3PEPS also provides a direct means to characterize conformational heterogeneity, as the asymptotic 3PEPS signals provide a quantitative measure of inhomogeneous broadening, which corresponds to the magnitude of the structural heterogeneity of the chromophore's environment (27, 31).

We reported (30) that Ab 4-4-20 evolved to bind fluorescein (Fl) through the introduction of two $V_L$ mutations that are located in CDR1 and an adjacent $\beta$-strand, and 10 $V_H$ mutations that are dispersed throughout the heavy-chain variable region (Fig. 1 and Fig. 5, which is published as supporting information on the PNAS web site). We prepared the chimeric Ab consisting of the germ-line light chain and mature heavy chain ($V_L^{gl}V_H^{4-4-20}$) and used surface plasmon resonance and 3PEPS to demonstrate that the two $V_L$ mutations increase affinity and shift $\rho_{Ab}(\omega)$ to higher frequency (30). We now report production and characterization of the full germ-line Ab ($V_L^{gl}V_H^{gl}$) and all of the heavy-chain single-point mutants. Using surface plasmon resonance, 3PEPS, and DSS, we show that the 10 $V_H$ mutations introduced during Ab evolution significantly rigidify the combining site and restrict it to a single, well defined conformation. Furthermore, based on the crystal structure of the mature Ab ($V_L^{4-4-20}V_H^{4-4-20}$)–Fl complex, along with a computational model of the germ-line Ab–Fl complex, we show that this rigidification occurred through the introduction of mutations far from the active site that mediate hydrogen bonds and packing interactions that cross-link the $\beta$-strands of the Ab. This mechanism of evolution-mediated conformational restriction should help us understand how other Abs, and proteins in general, are evolved for function.

## Results and Discussion

The heavy chain of $V_L^{4-4-20}V_H^{4-4-20}$ evolved from its germ-line progenitor, $V_L^{gl}V_H^{gl}$, via the introduction of 10 mutations (30). Based on the crystal structure of $V_L^{4-4-20}V_H^{4-4-20}$ bound to Fl (34), it is apparent that, unlike the light-chain mutations characterized previously, none of these heavy-chain mutations are in direct contact with the bound Fl (Fig. 1). To determine whether these mutations contribute to Fl recognition, we prepared $V_L^{gl}V_H^{gl}$ and $V_L^{4-4-20}V_H^{4-4-20}$ variants where, one at a time, each of the 10 $V_H$ mutations was changed back to the corresponding germ-line residue. Binding of Fl was analyzed for the $V_H$ single mutants in the context of both a germ-line and a mature $V_L$ (Table 1; $V_L^{gl}V_H^{P17S}$ and $V_L^{gl}V_H^{R38C}$ were not characterized because of poor expression). We observed that the dissociation constant, $K_D$, the dissociation rate constant, $k_{off}$, and the association rate constant, $k_{on}$, of Fl all strongly depend on the maturation state of the $V_L$. In the context of the germ-line $V_L$, all but one of the $V_H$ mutants bound Fl more tightly with the mature residue, as expected,

**Table 1. Binding affinity and kinetic data**

| Mutant | $K_D$, $\mu$M | $k_{on}$, $10^3$ $M^{-1}\cdot s^{-1}$ | $k_{off}$, $10^{-3}$ $s^{-1}$ | Mutant | $K_D$, $\mu$M | $k_{on}$, $10^3$ $M^{-1}\cdot s^{-1}$ | $k_{off}$, $10^{-3}$ $s^{-1}$ |
|---|---|---|---|---|---|---|---|
| $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ | 2.6 ± 0.6 | 16 ± 1 | 43 ± 3 | $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ | 0.22 ± 0.05 | 31 ± 12 | 7 ± 1 |
| $V_L{}^{gl}V_H^{P17S}$ | — | — | — | $V_L^{4\text{-}4\text{-}20}V_H^{P17S}$ | 0.18 ± 0.09 | 16 ± 8 | 2.9 ± 0.4 |
| $V_L^{gl}V_H^{S30T}$ | 3.9 ± 0.7 | 0.2 ± 0.1 | 0.9 ± 0.1 | $V_L^{4\text{-}4\text{-}20}V_H^{S30T}$ | 0.25 ± 0.06 | 3.7 ± 2 | 0.9 ± 0.4 |
| $V_L^{gl}V_H^{D31N}$ | 1.8 ± 0.4 | 2 ± 1 | 3 ± 2 | $V_L^{4\text{-}4\text{-}20}V_H^{D31N}$ | 0.03 ± 0.02 | 92 ± 6 | 2.3 ± 0.7 |
| $V_L^{gl}V_H^{Y32S}$ | 150 ± 28 | 0.2 ± 0.1 | 30 ± 0.2 | $V_L^{4\text{-}4\text{-}20}V_H^{Y32S}$ | 0.7 ± 0.2 | 4 ± 4 | 3 ± 3 |
| $V_L^{gl}V_H^{V37F}$ | 17 ± 10 | 0.3 ± 0.1 | 5 ± 1 | $V_L^{4\text{-}4\text{-}20}V_H^{S37F}$ | 0.06 ± 0.01 | 77 ± 25 | 4.6 ± 0.6 |
| $V_L^{gl}V_H^{R38C}$ | — | — | — | $V_L^{4\text{-}4\text{-}20}V_H^{R38C}$ | 1.3 ± 0.2 | 2 ± 2 | 2.6 ± 0.9 |
| $V_L^{gl}V_H^{R52K}$ | 9 ± 1 | 0.5 ± 0.1 | 4.1 ± 0.4 | $V_L^{4\text{-}4\text{-}20}V_H^{R52K}$ | 0.34 ± 0.07 | 10 ± 7 | 3.3 ± 0.4 |
| $V_L^{gl}V_H^{N52aS}$ | 7 ± 3 | 0.3 ± 0.1 | 2.4 ± 0.1 | $V_L^{4\text{-}4\text{-}20}V_H^{N52aS}$ | 0.14 ± 0.03 | 36 ± 10 | 5 ± 2 |
| $V_L^{gl}V_H^{V84A}$ | 9 ± 2 | 1.0 ± 0.6 | 8.6 ± 0.5 | $V_L^{4\text{-}4\text{-}20}V_H^{V84A}$ | 0.07 ± 0.05 | 46 ± 6 | 3.2 ± 0.5 |
| $V_L^{gl}V_H^{M87T}$ | 16 ± 7 | 0.3 ± 0.2 | 5.3 ± 0.1 | $V_L^{4\text{-}4\text{-}20}V_H^{M87T}$ | 0.09 ± 0.03 | 39 ± 12 | 3.3 ± 0.3 |

Binding affinity data for the $V_H$ single mutants in the context of a germ-line and mature $V_L$. Data for the $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ are shown for comparison.

because somatic mutations are only selected if they increase affinity. On average, the introduction of the mature residue increased the $K_D$ by 4.2-fold. However, in the context of the mature $V_L$, only 3 of the 10 $V_H$ mutants bound Fl more tightly with the mature residue, and by an average of only 2.2-fold, whereas 7 of the 10 $V_H$ mutants bound Fl more tightly with the corresponding germ-line residue, by an average of 3.2-fold. The $V_H$ mutations also had consistent effects on $k_{on}$ and $k_{off}$ in the context of a germ-line, but not a mature, $V_L$. The data demonstrate that at least the majority of the $V_H$ mutations must have been selected before the $V_L$ mutations, because they only increase affinity for Fl in the context of a germ-line $V_L$. Thus, $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ may be assumed to represent an evolutionary intermediate on the pathway that evolved the germ-line into the mature Ab. It is interesting to note that some of the most significant changes in binding kinetics and thermodynamics resulted from mutations at the residues that are most distant from Fl. This finding is consistent with the hypothesis that these mutations act indirectly, by altering protein structure or dynamics. Moreover, the effects of each mutation are clearly nonadditive, suggesting that the mutations act cooperatively to increase affinity for Fl.

To examine the contribution of the heavy-chain mutations to the evolution of Ab dynamics, we characterized $V_L^{gl}V_H^{gl}$ by using 3PEPS and DSS spectroscopy. We also reproduced the results of previous 3PEPS studies (30) of $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ and further characterized the two Abs by DSS spectroscopy. The 3PEPS and DSS experiments revealed that the mature $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ and its two evolutionary precursors are markedly different, both in terms of conformational heterogeneity and flexibility (Fig. 2). Most notably, we observed distinct differences in the asymptotic 3PEPS signals for each protein. $V_L^{gl}V_H^{gl}$ exhibits a large nonzero asymptote, which is significantly reduced in $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and virtually absent in $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ (Fig. 2*a*). The presence of the static inhomogeneity demonstrates that the germ-line Ab populates a broad distribution of different combining-site conformations that do not interconvert on the time scale of the experiment (0.3 ns). Conversely, the absence of static inhomogeneity in the mature Ab suggests that it binds Fl in a relatively well defined conformation.

The conclusion that evolution restricted the conformational heterogeneity of the Ab is supported by fluorescence lifetime measurements (Fig. 3). Only $V_L^{gl}V_H^{gl}$ shows a broad lifetime distribution with multiple maxima, further demonstrating that Fl is bound to a distribution of distinct combining-site conformations that do not interconvert on the fluorescence time scale (ns). The lifetime distribution narrows for $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and exhibits only a single peak for $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$, suggesting that when bound to the mature Ab, Fl experiences a single, well ordered binding site. In all, the 3PEPS and fluorescence lifetime experiments provide convincing evidence that evolution, especially of the heavy chain, acts to reduce the structural heterogeneity of the combining site by localizing it to a single conformation.

Interestingly, the 3PEPS and DSS signals also reveal a correlation between conformational heterogeneity and the amplitude of low-frequency protein motion. This correlation is apparent from a comparison of the $\rho_{Ab}(\omega)$, which we calculated from combining the 3PEPS and DSS data (Fig. 1). Comparison of the $\rho_{Ab}(\omega)$ for $V_L^{gl}V_H^{gl}$ and $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ shows that $V_H$ evolution reduced the amplitude of the ns time-scale motion and increased the amplitude of the ps time-scale motion. Comparison of $\rho_{Ab}(\omega)$ for

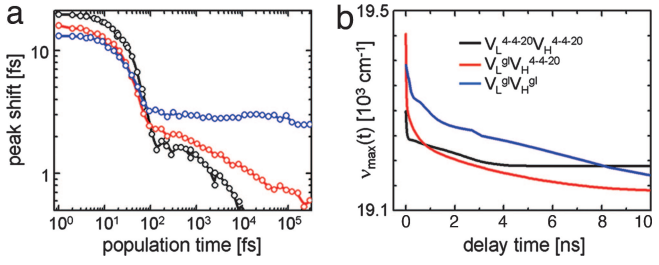


**Fig. 2.** Dynamic properties of the germ-line Ab $V_L^{gl}V_H^{gl}$ (blue), the intermediate $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ (red), and the mature Ab $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ (black). (*a*) 3PEPS decay, which measures the rephasing capability of the ensemble that decays because of protein fluctuations (27). The longtime signal offset in $V_L^{gl}V_H^{gl}$ indicates a large structural heterogeneity that is reduced in $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and virtually absent in $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$. (*b*) Time-dependent spectral position of the fluorescence maximum that shifts because of protein relaxation after photoexcitation of Fl (32, 33).

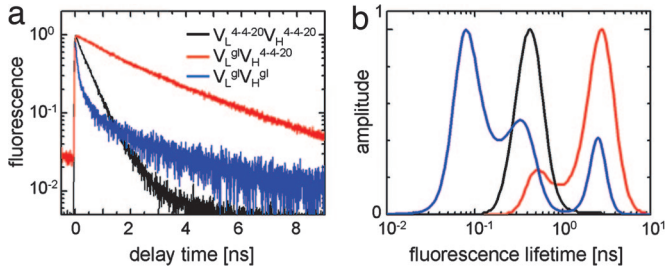


**Fig. 3.** Fluorescence decay of Fl bound to the germ-line Ab $V_L^{gl}V_H^{gl}$ (blue), the intermediate $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ (red), and the mature Ab $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ (black). (*a*) Magic angle fluorescence decay measured at 518 nm. (*b*) Fluorescence lifetime distributions obtained with the maximum-entropy method (41).

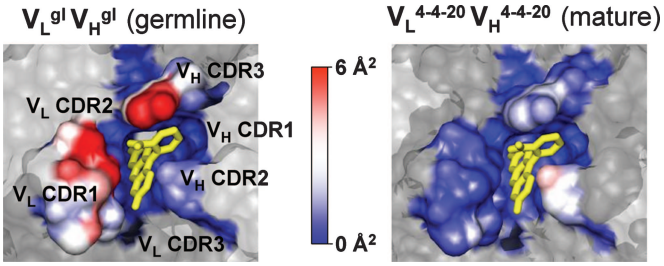


**Fig. 4.** rmsd values of residues in the Ab-combining site with bound Fl (yellow) for germ-line and mature Abs, obtained from classical MD simulations based on the crystal structure of the mature Ab (Protein Data Bank ID code 1FLR; ref. 34) and a computational model of the germ-line Ab.

$V_L^{gl}V_H^{4\text{-}4\text{-}20}$ and $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ shows that $V_L$ evolution further reduced both the ns and ps time-scale motions, resulting in a mature Ab that responds to Fl excitation with almost exclusively high-frequency motion. The correlation between conformational heterogeneity and low-frequency protein motion suggests that evolution localized the combining to a single conformation by significantly restricting protein motions.

The conclusion that evolution restricted the heterogeneity of the Ab-combining site by significantly restricting protein motions is further supported by molecular dynamics (MD) simulations. Starting from the crystal structure of $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ (34) or a computational model of $V_L^{gl}V_H^{gl}$, we ran 10-ns trajectories and calculated rmsd values, which reflect the mean displacement of atoms about the average structure (Fig. 4). We observed a significant decrease in the rmsd values for residues throughout the $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ combining site, relative to the $V_L^{gl}V_H^{gl}$ combining site. The changes were most pronounced in $V_L$ CDR1 and $V_H$ CDR3, which suggests that the mutations introduced during evolution rigidify the combining site by restricting the motion of the CDR loops.

We also used the trajectories to calculate the contributions of individual residue motions to $\rho_{Ab}(\omega)$. We found that in both the germ-line and mature Abs virtually all of the residues in the combining site contribute to $\rho_{Ab}(\omega)$, at least moderately, and no single residue contributes >15% of the total amplitude (Table 3, which is published as supporting information on the PNAS web site). When comparing the motions of different residues within a given combining site, similar time scales were observed, suggesting that the motions are correlated. However, for a given residue, the time scales are significantly shorter in the mature Abs, whereas the contribution to the total amplitude of $\rho_{Ab}(\omega)$ remains nearly constant (Table 3). Thus, the MD simulations suggest that the $\rho_{Ab}(\omega)$ reflect collective motions of the entire combining site and that the shift of $\rho_{Ab}(\omega)$ to higher frequency results from reduced CDR loop motion.

To more quantitatively interpret the evolution-induced rigidification and place it in the context of other proteins, we defined an average harmonic force constant of the combining site, $\langle k\rangle = \langle m\rangle\cdot\langle\omega^2\rangle$, where $\langle m\rangle$ is an average mass, and $\langle\omega^2\rangle$ may be obtained from the normalized second moment of $\rho_{Ab}(\omega)$

$$\langle\omega^2\rangle = \int \omega^2\rho_{Ab}(\omega)d\omega \Big/ \int \rho_{Ab}(\omega)d\omega.$$

$\langle\omega^2\rangle$ does not depend on the absolute value of $\rho_{Ab}(\omega)$, which allows the use of the $\rho_{Ab}(\omega)$ obtained from 3PEPS and DSS, even if they are not in units of rmsd (see *Supporting Text*, which is published as supporting information on the PNAS web site). Because the experimentally observed energy gap fluctuations are caused by side-chain and backbone atom fluctuations associated with collective protein motions, we assume a lower limit for $\langle m\rangle$ of 100 g/mol (the approximate mass of a single residue). Correspondingly, $\langle k\rangle$ values of $1\times10^{-3}$ N/m, $5\times10^{-3}$ N/m, and 0.4 N/m are obtained for the germ-line, intermediate, and mature proteins, respectively. Although the force constants are only approximate, they are likely to accurately reflect the relative changes induced by evolution, because the MD simulations predict similar reduced masses for the motion observed in each Ab (Table 3). Thus, we conclude that the Ab–Fl complex was systematically rigidified during evolution, ultimately increasing its rigidity ≈400-fold. Interestingly, the average force constant for the mature Ab is similar to the force constant of diffusive motions in myoglobin (0.3 N/m) measured in temperature-dependent neutron-scattering experiments (26), suggesting that the level of rigidity selected during Ab evolution is similar to that selected in other functional proteins.

A physical basis for the cooperative rigidification induced by the $V_H$ mutations is apparent from a comparison of the structure of the mature Ab (34) with the computational structure of the germ-line Ab. Five of the $V_H$ mutants are located in a proximal cluster that is ≈10 Å from the combining site, whereas the remaining five are located in a cluster that is >20 Å removed from the combining site (Fig. 1 and Fig. 6, which is published as supporting information on the PNAS web site). The cluster of mutations proximal to the combining site introduces interactions that cross-link all three $V_H$ CDR loops. Specifically, $S^H52aN$ introduces an H bond between CDR2 and the $D^H31$ backbone carbonyl of CDR1 that was also introduced by somatic mutation ($N^H31D$). This interaction may be stabilized by two adjacent somatic mutations, $T^H30S$ and $S^H32Y$. In addition, $S^H32Y$ introduces packing interactions with $Y^H100d$ and $Y^H102$ of $V_H$ CDR3. Packing interactions are also introduced between $V_H$ CDR1 and CDR2 by $K^H52R$, which packs on the side chain of $W^H33$. In a similar manner, mutations in the distal cluster introduce H bonds that cross-link the β-strands connected to $V_H$ CDR1 and CDR3. Central to these distal cluster interactions is the $C^H38R$ mutation, which appears to nucleate an H-bonding network involving $D^H86$, $Y^H90$, and $R^H66$. The mutations $A^H84V$ and $T^H87M$ may help stabilize these interactions. Overall, the $V_H$ mutations appear to rigidify the protein by introducing two clusters of mutually dependent interactions that act to cross-link β-strands and CDR loops of the combining site.

The experimental and simulation data may be combined to generate a picture of how evolution tailored the energy landscape of the Ab–Fl complex to restrict its conformational heterogeneity (Fig. 1). The immune system first responded to Fl with a germ-line Ab that populates different and relatively flexible conformations. $V_H$ mutations were then selected that introduced H-bonding and packing interactions that cross-link the loops and β-strands that form and support the combining site, respectively. This process resulted in a significant rigidification of the combining site, which increased the barrier to interconversion with other conformations. Finally, the two $V_L$ mutations, $H^L34R$ and $L^L46V$, introduced and optimized an H bond between the protein and Fl (30). Thus, during evolution an appropriate combining site was first selected from an ensemble of conformations populated by the flexible germ-line receptor and then the selected combining site was further optimized for recognition of the target molecule. A similar mechanism may also contribute to the evolution of other proteins, where mutations are suggested to have converted flexible, polyspecific, or functionally "promiscuous" proteins into more rigid and specific proteins (2, 35, 36).

## Materials and Methods

All Abs were expressed as Fab fragments (29, 30, 37). After isolation from the cell lysates by protein G affinity chromatography, Ab Fab fragments were further purified by cation ex-

change chromatography (Mono S; Amersham Pharmacia, Piscataway, NJ).

The dissociation constant $K_D$, the dissociation rate constant $k_{off}$, and the association rate constant $k_{on}$ of the Ab–Fl complexes were determined by using surface plasmon resonance on a Biacore 3000 biosensor (Biacore, Uppsala, Sweden) following published methods (38). Briefly, BSA was conjugated with Fl and immobilized on a research-grade CM5 sensor chip. $K_D$ was measured under equilibrium conditions, and $k_{off}$ was measured under kinetic conditions. The association rate constant was calculated as $k_{on} = k_{off}/K_D$.

The experimental setup for 3PEPS experiments has been described (28). In brief, samples were excited at 498–510 nm with 50 fs, 5- to 10-nJ pulses at 5-kHz repetition rate. Samples typically contained 100 $\mu$M Ab and 80 $\mu$M Fl in 10 mM Tris·HCl, pH 7.5. A spinning cell with a path length of 0.25 mm was used and maintained at 22 ± 1°C. 3PEPS signals in two phase-matching directions were detected with large-area avalanche photodiodes (Advanced Photonics, Irvine, CA).

Fluorescence kinetics were measured by a time-correlated single photon counting (TCSPC) setup as described (39). In brief, samples were excited at 464 or 488 nm with 0.3-nJ pulses at an 83-MHz repetition rate polarized at a magic angle with respect to a Glan-Thomson polarizer in the emission path. The instrument response function measured with scattered excitation light was typically 30 ps. Samples contained 100–700 $\mu$M Ab and ≈30 $\mu$M Fl in 10 mM Tris·HCl, pH 7.5 and were purged with Ar for 30 min before the experiment. Samples were stirred continuously in a 1-mm quartz cuvette. A 505DRLP dichroic filter (Omega, Brattleboro, VT) was used to block scattered excitation light. Fluorescence was detected by an R3809U-50 MCP (Hamamatsu, Middlesex, NJ) and an SPC-630 TCSPC module (Becker & Hickl, Berlin, Germany) through a 270M dual-port monochromator (Spex, Edison, NJ).

DSS data were obtained from fluorescence decays at 24 wavelengths with 50-cm$^{-1}$ spacing. The data sets were fit to the convolution of the instrument response function with a model function composed of a sum of exponentials, a baseline, and a time offset. The time-dependent fluorescence spectra were reconstructed by normalizing the integrated intensity from the deconvoluted kinetics probed at each wavelength to the steady-state fluorescence spectrum. These reconstructed spectra were then fit to log-normal functions to determine the spectral maximum, $\omega_{eg}(t)$ (40). Because of low affinity, $V_L^{gl}V_H^{gl}$ samples contained ≈8% of unbound Fl. To correct for signal from unbound Fl, we deconvoluted the time-dependent fluorescence spectra for each delay time into three Gaussian bands, one for the unbound dye, one for the bound dye, and one for the vibronic band on the red side of the spectrum (see Fig. 7, which is published as supporting information on the PNAS web site). The spectral position of the Gaussian that accounted for the free dye did not change after an initial fast decay (as observed for Fl in buffer). The Gaussians that accounted for the protein-bound dye and the vibronic band red-shifted with increasing delay time.

Fluorescence lifetime distributions were determined from magic angle fluorescence decays with the program MEMexp 3.0 developed by Peter Steinbach, National Institutes of Health, Bethesda, MD (41). To account for unbound dye in the $V_L^{gl}V_H^{gl}$ sample, a monoexponential decay with the lifetime of the unbound dye (4.7 ns) was convoluted with the instrument response function and subtracted from the experimental fluorescence decay of $V_L^{gl}V_H^{gl}$ before calculating the fluorescence lifetime distribution.

The 3PEPS decays were used to determine the high-frequency part (>0.5 cm$^{-1}$ corresponding to protein dynamics faster than 100 ps) of $\rho_{Ab}(\omega)$, and the static inhomogeneity, $\Delta_{in}$, as described (29, 30). Briefly, the experimental 3PEPS decay was fit by a model spectral density by using Mukamel's response function formalism (31). The total spectral density, $\rho(\omega)$, is the sum of both the intramolecular vibrations of the chromophore, $\rho_{Fl}(\omega)$, and the protein, $\rho_{Ab}(\omega)$. Intramolecular vibrational frequencies and excitation-induced displacements of Fl were obtained from quantum chemical calculations (see supporting information of ref. 29). $\rho_{Ab}(\omega)$ was modeled as the sum of two components: $\rho_{Ab}(\omega) = \rho_{BO}(\omega) + \rho_K(\omega)$. A Brownian oscillator term

$$\rho_{BO}(\omega) = \frac{2}{\pi\omega}\frac{\lambda_{BO}\omega_{BO}\Gamma_{BO}}{(\omega_{BO}{}^2 - \omega^2)^2 + \Gamma_{BO}{}^2\omega^4} \quad [1]$$

was used to represent the inertial sub-ps protein dynamics, where $\lambda_{BO}$ is the reorganization energy (corresponding to the amplitude of motion), $\omega_{BO}$ is the frequency, and $\Gamma_{BO}$ is the damping constant of the Brownian oscillator (31). Because amplitude $\lambda_{BO}$, frequency $\omega_{BO}$, and damping constant $\Gamma_{BO}$ of the fastest motion, corresponding to the ≈100-cm$^{-1}$ peak in $\rho_{Ab}(\omega)$, could not be fit unambiguously, we assumed that $\lambda_{BO}$ and $\omega_{BO}$ were identical in each Ab, and only the damping constant $\Gamma_{BO}$ was varied. This approximation does not affect the conclusion that the observed changes in $\rho_{Ab}(\omega)$ reflect Ab rigidification (30). In addition, because of the rather different time scales (separated by at least one order of magnitude), the parameters used to fit the ps and ns dynamics were independent of the specific model for the sub-ps dynamics.

A sum of Lorentzian terms according to overdamped Brownian oscillators was used to represent the ps dynamics

$$\rho_K(\omega) = \frac{1}{\pi\omega}\sum_i \frac{\lambda_{K,i}\tau_{K,i}}{1 + (\omega\tau_{K,i})^2}, \quad [2]$$

where $\lambda_{K,i}$ and $\tau_{K,i}$ are the reorganization energy and time constant of the $i^{th}$ mode, respectively.

Signals for the various time-resolved experiments such as 3PEPS and DSS and the steady-state absorption and emission spectra may be calculated from the line-broadening function $g(t)$ by using standard procedures (31). $g(t)$ may be calculated from $\rho(\omega)$ by using the expression

$$g(t) = \int_0^\infty \rho(\omega)\coth(\hbar\omega/2k_BT)(1 - \cos\omega t)d\omega + \Delta_{in}^2 t^2/2. \quad [3]$$

The parameters in $\rho_{Ab}(\omega)$ and the amount of static inhomogeneity ($\Delta_{in}$) in $g(t)$ were varied to obtain the best fit for the experimental data by using fit programs based on the program suite developed by Delmar Larsen, University of California, Davis. Fit results are listed in Table 3.

The low-frequency part of $\rho_{Ab}(\omega)$ (<0.5 cm$^{-1}$ corresponding to protein dynamics slower than 100 ps) is constructed by combining the results of 3PEPS and DSS experiments. We found ns kinetics in the DSS experiments (Table 3), but the conversion from DSS amplitudes into reorganization energies is not straightforward (31). Because the static inhomogeneity, $\lambda_{in}$, determined with 3PEPS gives an upper limit for the reorganization energy of low-frequency motion, we modeled the low-frequency part of $\rho_{Ab}(\omega)$ by using a Lorentzian (Eq. **2**) with an amplitude of $\lambda_{in}$ and a time constant determined from the DSS experiment. This approach was expected to accurately reproduce the frequency shifts and at least qualitatively reflect the relative amplitude changes for each Ab.

Table 2 lists the parameters used to fit the 3PEPS and DSS data, and the resulting $\rho_{Ab}(\omega)$ are shown in Fig. 1. It is interesting to note that while the amplitudes of the ps and ns dynamics ($\lambda_K$, $\lambda_{DSS}$) significantly vary between the three Abs, the corresponding time constants appear to be rather similar (3–5 ps and ≈3.5

**Table 2. Fit parameters for spectral density**

| Mutant | $\lambda_{BO}$, $cm^{-1}$ | $\omega_{BO}$, $cm^{-1}$ | $\Gamma_{BO}$, $cm^{-1}$ | $\lambda_K$, $cm^{-1}$ | $\tau_K$, ps | $\lambda_{inh}$, $cm^{-1}$* | $\lambda_{DSS}$, $cm^{-1}$† | $\tau_{DSS}$, ps† |
|---|---|---|---|---|---|---|---|---|
| $V_L^{gl}V_H^{gl}$ | 180 | 200 | 380 | | | 285 | 190 | 3,200 |
| $V_L^{gl}V_H^{4\text{-}4\text{-}20}$ | 180 | 200 | 580 | 50 | 3.2 | 121 | 93 | 3,700 |
| $V_L^{4\text{-}4\text{-}20}V_H^{4\text{-}4\text{-}20}$ | 180 | 200 | 620 | 20 | 5.0 | 6 | 27 | 500 |

Fit parameters used to fit 3PEPS and DSS data. For definition of parameters see Eqs. **1** and **2.**

*Static inhomogeneity, $\lambda_{in} = \Delta_{in}^2/2k_BT$.

†The DSS data were fit to a monoexponential decay: $I(t) = \lambda_{DSS}\cdot\exp(t/\tau_{DSS})$.

ns), suggesting that the effective masses of protein motions do not change significantly. At the same time, the amplitudes of the sub-ps motions are less affected by evolution, consistent with their interpretation as side-chain and small-scale motions inherent to any protein. The steady-state spectra for each Ab complex are shown in Fig. 8, which is published as supporting information on the PNAS web site, and the data are listed in Table 4, which is published as supporting information on the PNAS web site.

The computational model of the germ-line Ab–Fl complex was produced from the crystal structure of the Ab–Fl complex (Protein Data Bank ID code 1FLR; ref. 34) by changing the 12 somatic mutations by using the MMTSB tool set (42) and subjecting the structure to 1,000 steps of energy minimization by using the steepest-descent algorithm. This process was followed by another 300 minimization steps during which no coordinates were constrained. Using these structures, classical MD simulations using CHARMM (43) were performed in the canonical (NVT) ensemble at 298 K using 2-fs time steps in the velocity Verlet scheme (44) and constraining all bond distances between hydrogen and heavy atoms with the SHAKE algorithm (45). To reduce computation time, we removed the constant domains of the Fab fragment and used harmonic constraints of 1 kcal/mol per $Å^2$ to the regions further than 17 Å away from Fl to prevent unraveling of the variable domain. This system was solvated with a 21-Å sphere of TPI3 water (46) centered at Fl. Coordinates of Ab and Fl were held fixed for an equilibration period of 200 ps. MD trajectories of 10 ns were propagated, and vertical electronic transition energies were calculated by using snapshots of the trajectory by replacing the ground-state charge distribution of the chromophore with the excited-state charge distribution.

We thank Dr. Tomas Mancal and Prof. Minheang Cho for helpful discussions. This research was supported by The Skaggs Institute for Chemical Biology (F.E.R.).

1. Nagao C, Terada T, Yomo T, Sasai M (2005) *Proc Natl Acad Sci USA* 102:18950–18955.
2. O'Brien P, Herschlag D (1999) *Chem Biol* 6:R91–R105.
3. James L, Tawfik D (2003) *Trends Biochem Sci* 28:361–368.
4. Koshland DE, Jr (1958) *Proc Natl Acad Sci USA* 44:98–104.
5. Berzofsky JA (1985) *Science* 229:932–940.
6. Sundberg EJ, Mariuzza RA (2003) *Adv Protein Chem* 61:119–160.
7. Foote J, Milstein C (1994) *Proc Natl Acad Sci USA* 91:10370–10374.
8. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC (1997) *Science* 276:1665–1669.
9. Ma B, Shatsky M, Wolfson HJ, Nussinov R (2002) *Protein Sci* 11:184–197.
10. Comtesse N, Heckel D, Maldener E, Glass B, Meese E (2000) *Clin Exp Immunol* 121:430–436.
11. Souroujon M, White-Scharf ME, Andre-Schwartz J, Gefter ML, Schwartz RS (1988) *J Immunol* 140:4173–4179.
12. Jegerlehner A, Storni T, Lipowsky G, Schmid M, Pumpens P, Bachmann MF (2002) *Eur J Immunol* 32:3305–3314.
13. Baumgarth N (2000) *Immunol Rev* 176:171–180.
14. Guigou V, Guilbert B, Moinier D, Tonnelle C, Boubli L, Avrameas S, Fougereau M, Fumoux F (1991) *J Immunol* 146:1368–1374.
15. Mason D (1998) *Immunol Today* 19:395–404.
16. James LC, Roversi P, Tawfik DS (2003) *Science* 299:1362–1367.
17. Patten PA, Gray NS, Yang PL, Marks CB, Wedemayer GJ, Boniface JJ, Stevens RC, Schultz PG (1996) *Science* 271:1086–1091.
18. Joyce GF (1997) *Science* 276:1658–1659.
19. Hodgkin PD (1998) *Immunologist* 6:223–226.
20. Yin J, Beuscher A, Andryski S, Stevens R, Schultz P (2003) *J Mol Biol* 330:651–656.
21. Notkins A (2004) *Trends Immunol.* 25:174–179.
22. Manivel V, Sahoo N, Salunke D, Rao K (2000) *Immunity* 13:611–620.
23. Sagawa T, Oda M, Ishima M, Furukuwa K, Azuma T (2003) *Mol Immunol* 39:801–808.
24. Frauenfelder H, Sligar S, Wolynes P (1991) *Science* 254:1598–1603.
25. Parak F (2003) *Curr Opin Struct Biol* 13:552–557.
26. Zaccai G (2000) *Science* 288:1604–1607.
27. Fleming GR, Cho M (1996) *Annu Rev Phys Chem* 47:109–134.
28. Jimenez R, Case DA, Romesberg FE (2002) *J Phys Chem B* 106:1090–1103.
29. Jimenez R, Salazar G, Baldridge KK, Romesberg FE (2003) *Proc Natl Acad Sci USA* 100:92–97.
30. Jimenez R, Salazar G, Yin J, Joo T, Romesberg FE (2004) *Proc Natl Acad Sci USA* 101:3803–3808.
31. Mukamel S. (1995) *Principles of Nonlinear Optical Spectroscopy* (Oxford Univ Press, Oxford).
32. Jimenez R, Fleming GR, Kumar PV, Maroncelli M (1994) *Nature* 369:471–473.
33. Pierce DW, Boxer SG (1992) *J Phys Chem* 96:5560–5566.
34. Whitlow M, Howard AJ, Wood JF, Voss EW, Jr, Hardman KD (1995) *Protein Eng* 8:749–761.
35. Copley S (2003) *Curr Opin Chem Biol* 7:265–272.
36. Aharoni A, Leonid G, Khersonsky O, Gould SM, Roodveldt C, Tawfik D (2005) *Nat Genet* 37:73–76.
37. Ulrich HD, Patten PA, Yang PL, Romesberg FE, Schultz PG (1995) *Proc Natl Acad Sci USA* 92:11907–11911.
38. Yin J, Mundorff EC, Yang PL, Wendt KU, Hanway D, Stevens RC, Schultz PG (2001) *Biochemistry* 40:10764–10773.
39. McAnaney T, Shi X, Abbyad P, Jung H, Remington S, Boxer S (2005) *Biochemistry* 44:8701–8711.
40. Horng ML, Gardecki JA, Papazyan A, Maroncelli M (1995) *J Phys Chem* 99:17311–17337.
41. Steinbach P, Ionescu R, Matthews C (2002) *Biophys J* 82:2244–2255.
42. Feig M, Karanicolas J, Brooks C (2004) *J Mol Graphics Mod* 22:377–395.
43. MacKerell A, Brooks B, Brooks C, Nilsson L, Roux B, Won Y, Karplus M (1998) in *Encyclopedia of Computational Chemistry*, ed Schleyer P (Wiley, Chicester, UK), Vol 1, pp 271–277.
44. Swope WC, Andersen HC, Berens PH, Wilson KR (1982) *J Chem Phys* 76:637–649.
45. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) *J Comput Phys* 23:327–341.
46. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) *J Chem Phys* 79:926–935.