Published in final edited form as:

J Phys Chem B. 2010 June 3; 114(21): 7359–7370. doi:10.1021/jp906421v.

A Molecular Description of Flexibility in an Antibody Combining Site

Jörg Zimmermann^{*,†}, Floyd E. Romesberg[†], Charles L. Brooks III^{§,‡}, and Ian F. Thorpe^{§,¶,*}
Department of Molecular Biology and Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla California 92037

Abstract

Mature antibodies (Abs) that are exquisitely specific for virtually any foreign molecule may be produced by affinity maturation of naïve (or germline) Ab. However, the finite number of germline Ab available suggests that, in contrast to mature Ab, germline Ab must be broadly polyspecific so that they are able to recognize a wide range of ligands. Thus, affinity maturation must play a role mediating Ab specificity. One biophysical property that distinguishes polyspecificity from specificity is protein flexibility; a flexible combining site is able to adopt different conformations that recognize different foreign molecules (or antigens), while a rigid combining site is locked into a conformation that is specific for a given antigen. Recent studies (Proc. Natl. Acad. Sci. USA 104:8821–8826, 2007) have examined, at the atomic level, the structural properties that mediate changes in flexibility at four stages of affinity maturation in the 4-4-20 Ab. These studies employed molecular dynamics simulations to reveal a network of residue interactions that mediate the flexibility changes accompanying maturation. The flexibility of the Ab combining sites in these molecular systems was originally measured using 3-pulse photon echo spectroscopy (3PEPS). The present investigation extends this work by providing a concrete link between structural properties of the Ab molecules and features of the spectroscopic measurements used to characterize their flexibility. Results obtained from the simulations are in good qualitative agreement with the experimental measurements and indicate that the spectroscopic signal is sensitive to protein dynamics distributed throughout the entire combining site. Thus, the simulations provide a molecular level interpretation of the changes induced by affinity maturation of the Ab. The results suggest that 3PEPS spectroscopy in combination with molecular dynamics simulations can provide a detailed description of protein dynamics and, in this case, how it is evolved for biological function.

Keywords

3PEPS spectroscopy; molecular recognition; ligand association; binding; molecular dynamics

*Corresponding author. jzimm@scripps.edu, Phone : (858) 784-7335, Fax : (858) 784-7472. ithorpe@umbc.edu, Phone: (410) 455-5728, Fax: (410) 455-2608.

§Department of Molecular Biology

†Department of Chemistry

¶Current address: Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore MD, 21250

‡Current address: Department of Chemistry and Biophysics Program, University of Michigan, 930 North University Avenue, Ann Arbor, MI 48109

Supporting Information Available: Detailed information on the multi-exponential fits to the correlation functions and the covariance analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

1. Introduction

Antibodies (Abs) are the quintessential prototype for molecular recognition in biological systems, with the immune response generating Abs that selectively bind virtually any foreign molecule (or ‘antigen’). These mature Abs are evolved from a finite set of naïve (or ‘germline’) Abs by iterative cycles of somatic mutation and selection, which combined are known as affinity maturation. The fact that the germline Ab repertoire is finite (being limited by the number of B cell lymphocytes) suggests that, in contrast to mature Abs, germline Abs must be broadly polyspecific.¹ Polyspecificity would ensure that at least one member of the germline repertoire recognizes, at least with moderate affinity, any member of the virtually infinite set of antigens. Thus, affinity maturation is likely to play a role in generating Ab specificity.

A number of studies have been directed at establishing a link between physical properties of Abs and affinity maturation. The structural origins of affinity maturation have been elegantly studied by Schultz and Stevens, who characterized the structures of pairs of germline and mature Abs, both free and bound to their antigen.^{2–4} These studies revealed that, at least in some cases, affinity maturation pre-orders the Ab combining site to favor antigen binding. Li et. al. showed that maturation of Abs to hen egg white lysozyme was accompanied by burial of increased amounts of apolar surface and enhanced shape complementarity.⁵ Studies by Terzyan et. al. of a series of Abs that bind to the chromophore fluorescein (FL), including Ab 4-4-20 described in this work, suggest that structural changes in the combining site mediate the affinity increases that accompany maturation.⁶ Affinity maturation has also been shown to result in a more favorable entropy of antigen binding.^{7–11} These structural and thermodynamic data suggest that affinity maturation evolves polyspecific germline Abs in to more specific mature Abs, at least in part, by preordering the combining site in a conformation appropriate for antigen binding. This is consistent with the evolution of flexible Abs into more rigid receptors. A flexible combining site is able to adopt diverse conformations that recognize different antigens, while a rigid combining site is restricted to sample a more limited set of conformations, thus increasing specificity.

We are interested in more directly testing the hypothesis that affinity maturation tailors Ab dynamics and conformational heterogeneity. Toward this goal, we have employed Abs that were evolved to bind chromophoric antigens, such as FL^{12–15} and 8-methoxypyrene-1,3,6, trisulfonic acid (MPTS).¹⁶ We employed chromophoric antigens because they allow for the use of spectroscopic methods, such as three-pulse photon echo peak shift (3PEPS) spectroscopy,^{17–19} to characterize the protein that evolved to bind them^{12,16} yet are expected to be recognized like any other antigen. Our previous characterization of Ab 4-4-20 using 3PEPS spectroscopy revealed that the Ab is indeed significantly rigidified during affinity maturation.^{13–15} In general, 3PEPS experiments measure how the environment around a chromophore relaxes in response to a photoinduced change in the charge distribution of the chromophore. However, 3PEPS averages over all motions that are coupled to the chromophore’s transition dipole, thus it is not possible to interpret the data in terms of specific protein motions. While 3PEPS experiments thus lend direct support to the hypothesis that polyspecific germline Abs are evolved into more specific mature Abs by tailoring protein dynamics, the detailed mechanism by which this occurred for 4-4-20 remains to be fully understood. Because the observables in a 3PEPS experiments are related to protein dynamics via fluctuations in the energy gap between ground and excited states of the chromophore, their origins may be investigated using molecular dynamics (MD) simulations. Thus, a combination of 3PEPS experiments and MD simulations promises to provide a detailed interpretation of Ab dynamics and how they are evolved during affinity maturation.

The antigen binding or ‘combining’ site of an Ab is located at the interface of two polypeptides called light chain and heavy chain. All Ab sequences are very similar except in the N-terminal

domain of the light and heavy chains, which contain the ‘variable regions’, V_L in the light chain and V_H in the heavy chain (see Fig. 1). Within V_L and V_H there are six specific loops of hypervariable sequence, called the complementarity determining regions (CDRs), which form the combining site of the Ab. The light and the heavy chain each contribute three CDRs to the combining site (V_L CDR1-3 and V_H CDR1-3), which are supported by the so-called framework regions (V_L FR1-4 and V_H FR1-4). Previously we reported the sequence of mature 4-4-20 ($V_L^{4-4-20}V_H^{4-4-20}$) as well as its germline precursor ($V_L^{gl}V_H^{gl}$).^{14,20} In this nomenclature, V_L and V_H refer to the light or heavy chain of the Ab molecule, respectively, while the superscript refers to the protein maturation state: 4-4-20 denotes the mature Ab while *gl* denotes the germline Ab. By characterizing the affinity for FL as a function of converting each somatic mutation in 4-4-20 back to its germline residue, we identified two approximate intermediates ($V_L^{gl}V_H^{4-4-20}$ and $V_L^{H34R}V_H^{4-4-20}$) along the evolutionary pathway between the germline and mature Ab.¹³ The conversion of $V_L^{gl}V_H^{gl}$ to $V_L^{gl}V_H^{4-4-20}$ is associated with ten amino acid substitutions in the heavy chain, conversion of $V_L^{gl}V_H^{4-4-20}$ to $V_L^{H34R}V_H^{4-4-20}$ is associated with mutation of His^L34 (*i.e.* His34 of V_L) to Arg, and conversion of $V_L^{H34R}V_H^{4-4-20}$ to $V_L^{4-4-20}V_H^{4-4-20}$ is associated with mutation of Leu^L46 to Val.

Previously, MD simulations have been employed to understand the changes in flexibility which accompany the affinity maturation of 4-4-20.²⁰ These studies suggested that a network of interacting residues mediate the rigidification induced by affinity maturation. (It should be noted that in those studies a different nomenclature for the Ab molecules was employed:

$V_L^{gl}V_H^{gl}$, $V_L^{gl}V_H^{4-4-20}$, $V_L^{H34R}V_H^{4-4-20}$ and $V_L^{4-4-20}V_H^{4-4-20}$ respectively correspond to the GL, IM1, IM2 and AM designations used in that work). Here, we build on these results by elucidating the link between the structural properties of the Ab and features of the spectroscopic measurements used for their characterization. We employ MD simulations described above²⁰ to compute correlation functions for the transition energy in each of the Ab-FL complexes. The results show good qualitative agreement with experimental findings regarding the differing time scales of fluctuations that modulate the FL environment and indicate that the 3PEPS signals reflect motions of the entire combining site. A covariance analysis further shows that Ab fluctuations in both CDR's and framework regions are correlated, and that this correlation is stronger in the mature than in the germline Ab. These results provide additional evidence that the Ab is rigidified during maturation. We anticipate that this information will prove useful in the implementation and analysis of similar studies of protein dynamics as well as provide insight into the determinants of molecular recognition and how it is evolved for biological function. In the subsequent section we provide the theoretical background for our studies; this is followed by a description of the simulation methodology. Our observations and a discussion of their implications are then presented. We conclude with a summary of our findings.

2. Theory

2.1. Link between 3PEPS Signal and Transition Energy Fluctuations

3PEPS experiments measure the position of the signal maximum of the time-integrated photon echo (the ‘peak shift’) emitted after applying a sequence of three resonant laser pulses to the sample (Fig. 2A,C).¹⁷ The first pulse creates a coherent superposition between the ground and excited states of the chromophores, and the resulting ensemble of photoexcited chromophores subsequently dephases (Fig. 2B). The second pulse creates a population, either in the ground or excited state, preventing further dephasing. The third pulse again creates a superposition of states, allowing the ensemble to rephase and emit an ‘echo’ signal. Molecules may also relax via a free-induction decay (FID) process, which does not require rephasing.²¹ The intensity of the time-integrated signal originates both from those molecules that rephase and those that relax via FID. Environmental fluctuations (vibrations of the chromophore and vibrations and

diffusive motions of the protein) that occur during the time between interaction with the first and third laser pulse induce random phase changes, which prevent rephasing of the affected chromophores but do not affect the FID signal. Consequently, with longer delay time, proportionally more molecules relax by FID, causing an apparent shift of the peak position of the time-integrated signal toward zero (*i.e.* a decay of the peak shift, Fig. 2C).

Fleming,^{18,19} and Wiersma²² have established that in the high-temperature limit and for times longer than the bath-correlation time, the normalized 3PEPS decay approximates the two-point time correlation function of fluctuations in the electronic transition energy between ground and excited state, $\varepsilon(t) = H\omega_{eg}(t)$,

$$C(\tau) = \frac{\langle \delta\varepsilon(t)\delta\varepsilon(t+\tau) \rangle}{\langle \delta\varepsilon(t)^2 \rangle} \quad (1)$$

where $\delta\varepsilon(t) = \varepsilon(t) - \langle \varepsilon \rangle$ are the fluctuations in $\varepsilon(t)$ around its equilibrium, ensemble-averaged value, $\langle \varepsilon \rangle$, and $\tau \geq 0$. Note that for chromophore or solvent responses on a timescale similar to the bath correlation time, as is the case for the Ab-FL complexes, $C(\tau)$ can be determined by means of a more complex data evaluation of the 3PEPS decay based on linear response theory;^{17,18,21} this approach was employed to generate the experimental $C(\tau)$ listed in Table 1 from the 3PEPS signals.^{13,14} This procedure also allows the use of molecular simulation data to recreate the actual peak shift itself.^{23–25} However, molecular simulations can also be used to delineate the impact of protein fluctuations on the chromophore environment via direct analysis of $C(\tau)$.^{26–28} The time scales of protein motions that cause phase shifts of the transition dipole also govern the decay of $C(\tau)$.^{29,30} However, the $C(\tau)$ determined from a 3PEPS experiment averages over all protein motions weighted by their coupling strength to the chromophore's transition dipole, which is a function of distance and charge of the moving entity. Computational methods allow for the deconvolution of $C(\tau)$ into the contributions of individual protein motions and to account for the electrostatic weights, as described in the following section. Similar deconvolution of $C(\tau)$ into distinct contributions arising from protein or solvent fluctuations has been carried out for the proteins myoglobin^{26,31,32} and monellin.²⁷

2.2. Extracting Contributions to $C(\tau)$ from Specific Protein Groups

To identify specific interactions that contribute to the decay of $C(\tau)$ we make use of the fact that the molecular mechanics force field used to evaluate $\varepsilon(t)$ is strictly pairwise additive. Because there are no explicit multibody terms in the force field, $\varepsilon(t)$ and $\delta\varepsilon(t)$ can be modeled as a sum of contributions from the energies $\varepsilon_i(t)$ due to each atom i in the system,

$$\varepsilon(t) = \sum_i \varepsilon_i(t) \quad \text{and} \quad \delta\varepsilon(t) = \sum_i \delta\varepsilon_i(t), \quad \text{with } \delta\varepsilon_i(t) = \varepsilon_i(t) - \langle \varepsilon_i(t) \rangle.$$

$C(\tau)$ can thus be written as:

$$C(\tau) = \sum_i^n \frac{\langle \delta\varepsilon_i(t)\delta\varepsilon_i(t+\tau) \rangle}{\left\langle \left(\sum_i \delta\varepsilon_i(t) \right)^2 \right\rangle} + \sum_{i \neq j}^n \frac{\langle \delta\varepsilon_i(t)\delta\varepsilon_j(t+\tau) \rangle}{\left\langle \left(\sum_i \delta\varepsilon_i(t) \right)^2 \right\rangle} \quad (2)$$

where the summation runs over all n atoms in the system. For computational expediency, we carried out this analysis on a per residue basis rather than for individual atoms; i in the equations below thus will denote individual residues rather than atoms. Using the approach of Nilsson

and Halle²⁷, one can define a partial time correlation function, $C_i^{partial}(\tau)$ to assess the contribution that an individual residue makes to the time dependence of $C(\tau)$. In these studies we evaluate $C_i^{partial}(\tau)$ according to:

$$C_i^{partial}(\tau) = \frac{\left\langle \delta\epsilon_i(t) \sum_j^n (\delta\epsilon_j(t+\tau)) \right\rangle}{\left\langle \delta\epsilon_i^2(t) \right\rangle^{1/2} \left\langle \left(\sum_i^n \delta\epsilon_i(t) \right)^2 \right\rangle^{1/2}} \quad (3)$$

One can rank the contribution each $C_i^{partial}(\tau)$ makes $C(\tau)$ by normalizing this quantity via

$$a_i = \frac{\left\langle \delta\epsilon_i^2(t) \right\rangle^{1/2}}{\left\langle \left(\sum_i^n \delta\epsilon_i(t) \right)^2 \right\rangle^{1/2}} \quad a_i C_i^{partial}(\tau) \text{ with } \left\langle \left(\sum_i^n \delta\epsilon_i(t) \right)^2 \right\rangle^{1/2}. \text{ With this choice } C(\tau) \text{ can be recovered via}$$

$$C(\tau) = \sum_i^n a_i C_i^{partial}(\tau). \text{ Thus, the relative contribution of each residue to } C(\tau) \text{ is governed by } \delta\epsilon_i. \text{ The magnitude of } \delta\epsilon_i \text{ is a measure of the spatial amplitude of single-residue fluctuations weighted by the electrostatic coupling of the residue to the FL transition dipole.}$$

In an attempt to understand the molecular origin of the dephasing correlation function, it is reasonable to ask whether $C(\tau)$ can be expressed in terms of time scales which arise primarily from the fluctuations of individual protein groups. One can employ Eq. 2 to probe this question further. The first term in Eq. 2 contains only contributions from self-terms while the second term contains the contributions from cross-correlations. By modifying the normalization factors employed above, we can construct a quantity we define as the additive correlation function, $C(\tau)_{ad}$:

$$C(\tau)_{ad} = \sum_i^n a_i C_i(\tau) \quad (4)$$

$$a_i = \frac{\left\langle \delta\epsilon_i^2 \right\rangle}{\sum_i^n \left\langle \delta\epsilon_i^2 \right\rangle}; C_i(\tau) = \frac{\left\langle \delta\epsilon_i(t) \delta\epsilon_i(t+\tau) \right\rangle}{\left\langle \delta\epsilon_i^2 \right\rangle}.$$

with

$C(\tau)_{ad}$ is composed of the time-correlation functions $C_i(\tau)$ of individual residues. One can thus isolate the impact that fluctuations of a given residue have on the time scales present in $C(\tau)$ by computing $C(\tau)_{ad}$ to exclude cross-correlations and evaluating each $C_i(\tau)$ and a_i . The constant a_i determines the relative weight of an individual correlation function $C_i(\tau)$ in $C(\tau)_{ad}$. The values of each a_i are related to the variance s_i^2 for fluctuations in ϵ_i via:

$$a_i = \frac{\left\langle \delta\epsilon_i^2 \right\rangle}{\sum_i^n \left\langle \delta\epsilon_i^2 \right\rangle} = \frac{s_i^2}{\sum_i^n s_i^2} \quad (5)$$

Note that the time dependence of $\delta\epsilon_i$ no longer needs to be explicitly considered as the ensemble averages are, in principle, independent of time. Since the value of the denominator above is the same for every atom, it can be ignored when evaluating relative magnitudes. Therefore, the relative contribution of individual $C_i(\tau)$ terms in $C(\tau)_{ad}$ is completely described by computing s_i^2 . By extension, one can use s_i^2 to rank the contributions of individual residues to $C(\tau)_{ad}$. By neglecting cross-correlations, $C(\tau)_{ad}$ emphasizes the impact of fluctuations of individual residues. Thus, $C(\tau)_{ad}$ provides the optimal venue to test whether fluctuations of certain residues in the Ab may give rise to distinct decay time scales in $C(\tau)$.

3. Computational Methods

3.1. Molecular Simulations

Simulations were initiated from a crystal structure of the antigen binding fragment (Fab) of the mature 4-4-20 Ab (PDB designation 1FLR)³³ bound to FL. The charge distribution employed for the ground state of FL was determined by fitting point charges to the electrostatic potential of the molecule using the CHELPG fitting scheme.³⁴ Before fitting the electron density for point charges, the geometry of FL was optimized in vacuum at the B3LYP/6-31G+(d) level of theory using the Gaussian 98 program.³⁵ The resulting charges were scaled by 0.9 to be consistent with charge sets employed in the CHARMM molecular simulation program.^{36–38} Other parameters required for simulation of FL were obtained by assigning standard CHARMM atom types to FL atoms.

Structures of $V_L^{gl}V_H^{gl}$ and the two evolutionary intermediates were derived from the X-ray coordinates of $V_L^{4-4-20}V_H^{4-4-20}$ by performing appropriate amino acid substitutions using the Multiscale Modeling Tools for Structural Biology (MMTSB) Tool Set.³⁹ Coordinates of atoms not present in the crystal structure were assigned standard values from CHARMM topology files. The assigned coordinates as well as converted residues were subjected to 1000 steps of energy minimization using the steepest descent algorithm while keeping the remaining atoms fixed. This was followed by 300 minimization steps during which no coordinates were constrained. In preparing the simulations, emphasis was placed on regions of the Ab complexes in the vicinity of the combining site because the 3PEPS measurements probe the local environment of the FL chromophore.^{12,13,15} Thus, the Fab was truncated at the boundary between the constant and variable domains to make the simulations less computationally intensive (Fig. 1).

The combining site was solvated with a 21 Å sphere of TIP3P water⁴⁰ centered at the position of the center of mass of FL in the 1FLR crystal structure. Integrity of the solvent sphere was maintained through the use of a stochastic boundary potential.⁴¹ All bond distances between hydrogen and heavy atoms were constrained with the SHAKE algorithm.⁴² A spherical cutoff of 14 Å for non-bonded interactions was employed. The van der Waals interactions were gently switched to zero between 12–14 Å while forces corresponding to electrostatic interactions were shifted to zero at the cutoff distance. During an equilibration period of 200 ps the coordinates of the protein and ligand were restrained with mass weighted harmonic potentials which were gradually decreased from 5 to 1 kcal/mol Å². At the end of equilibration, these restraints were applied only to residues in the truncated region more than 17 Å from FL. The restraints were maintained throughout the simulations to maintain the integrity of the truncated Fab. MD trajectories were propagated for 10 ns in the canonical (NVT) ensemble using the Nosé-Hoover thermostat with a relaxation time of 1 ps. A time step of 2 fs was employed and coordinates written out every 100 steps. Equations of motion were integrated with the velocity Verlet scheme. Further simulation details were presented previously.²⁰ It should be noted that, in principle, application of a thermostat can alter the time scales of fluctuations in molecular dynamics simulations. However, it is known that for well equilibrated systems (such as those

employed in this work) the impact of the Nosé-Hoover thermostat on the fluctuations of any given particle is expected to be negligible.⁴³ Moreover, the thermostat has a similar impact on every particle in the system, allowing relative changes in timescales of protein motions to be correctly reproduced.

3.2. Evaluation of Correlation Functions

Explicit water molecules were removed from the MD trajectories and replaced with Generalized Born (GB) implicit solvent.⁴⁴ In principle, implicit solvent models such as GB provide the potential of mean force for the energetic contributions of solvent with respect to any given solute configuration, properly incorporating bulk solvent effects in a mean field sense. The use of an implicit solvent model in this context allows for contributions from water to be incorporated into ε in a mean field manner while minimizing the impact of solvent fluctuations on $\delta\varepsilon$. These considerations are particularly important in the present study because the sphere of water used to solvate the combining site is much smaller than a bulk water system (see Fig. 1) and is therefore subject to much larger energetic fluctuations. This procedure reduces the noise in $\delta\varepsilon$ derived from solvent fluctuations and allows us to focus on contributions made by protein residues.

Thus, the total energy ε is given by:

$$\varepsilon = \varepsilon_{\text{solute}} + \Delta G_{\text{solv}} \quad (5)$$

where the first term on the right is the molecular mechanics energy of the protein and ligand while ΔG_{solv} is the electrostatic free energy of solvation provided by the GB approach. A detailed description of the expressions used to evaluate ΔG_{solv} has been provided in previous work.⁴⁴

For each snapshot of these trajectories, a vertical excitation was performed by replacing the ground state charge distribution of FL with the corresponding charges for the excited state. Thus, $\varepsilon(t)$ was modeled by the total energy difference between the ground and excited state charge distributions within a given snapshot from the MD simulations. A similar approach was employed by Halder et al. in studies of apomyoglobin and apoleghemoglobin.²⁸ The excited state charge distribution of FL was determined by fitting point charges to the electrostatic potential obtained at the CIS/6-31G+(d) level of theory using the Gaussian 98 program³⁵ and the CHELPG fitting scheme.³⁴ As for the ground state (see above), excited state charges were uniformly scaled by 0.9. Ground and excited state charges for FL atoms are displayed in Fig. 3. The resulting time series for ε , $\varepsilon(t)$, were employed to compute $C(\tau)$ for each Ab as described in Section 2.1 (Fig. 4). Note that $C(\tau)$ contains contributions from all components of the simulation system and includes cross-correlations between the energetic contributions of individual components. The $C(\tau)$ for each Ab are shown in Fig 4 while time constants extracted from the $C(\tau)$ are displayed in Table 1.

The individual correlation functions described in Eqns. 3 and 4 were computed by removing all atoms except for those of FL and the atoms of interest from the MD snapshots before placing each snapshot in GB implicit solvent (Fig. 5). For computational expediency this procedure was carried out on a per residue basis rather than for individual atoms. In principle, the electrostatic coupling between the protein/solvent environment and the chromophore transition dipole depends on charge-dipole, dipole-dipole and higher order terms. Such interactions decay much more rapidly with distance than Coulombic (charge-charge) interactions. Thus, one would expect residues more than 7 Å away from FL to exhibit negligible values of s_i^2 . Consequently, such residues were excluded from the analysis. Visual inspection was further employed to ensure that only residues clearly interacting with charged atoms in FL were included, reducing the number of residues considered to ten.

The various $C_i(\tau)$ were used to generate $C(\tau)_{ad}$, with the relative magnitudes of a_i obtained by computing the variance $\varepsilon_i(t)$ as described in Eqn. 5. For increased sensitivity, the contribution of FL atoms to $\varepsilon_i(t)$ was removed before computing s_i^2 for individual residues. The values of s_i^2 were used to rank the contributions of individual residues. Time constants extracted from $C_i^{partial}(\tau)$, $C_i(\tau)$ and $C(\tau)_{ad}$ are shown in Tables 2 and 3.

All correlation functions were fit to multi-exponential decays with the program MEMexp,⁴⁵ using the hybrid maximum entropy/nonlinear least square fitting mode to determine the number of exponentials required to fit a given correlation function.

3.3 Covariance Analysis

The covariance of atomic displacements provides an efficient way to compare variability of the fluctuations present in proteins and reveals structural correlations between different components of the Ab-FL complexes. The covariance of spatial displacements of atoms in each Ab system, μ_{ij} , were computed via $\mu = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle$, where r_i represents the displacement of atom i . These covariance values represent correlation coefficients for the fluctuations of individual atoms. In order to carry out this analysis, explicit solvent was first deleted and average structures for each Ab-FL complex were computed. Translation and rotation were removed from each trajectory by performing an RMS fit to heavy atoms of the average structure. Atomic fluctuations with respect to this average structure were determined for FL and C_α atoms, and these fluctuations were employed to generate a mass weighted covariance matrix⁴⁶ using the CORMAN module of CHARMM. C_α correlations are representative of the correlations of each residue as a whole. Covariance matrices for $V_L^{4-4-20} V_H^{4-4-20}$ and $V_L^{gl} V_H^{gl}$ are shown in Fig. 6. In order to provide further insight into the patterns of correlation occurring in each system, we define the metric summed squared covariance, SSC. For a given site i , the SSC with respect to structural subunit J is defined as $SSC_{i,J} = \sum_{j \in J, j \neq i} \mu_{ij}^2$, where j denotes a site within the subunit. SSC values compactly illustrate the extent of correlation between different regions of the Ab-FL complexes.

4. Results and Discussion

4.1. Dephasing Time Scales Exhibit the Experimentally Observed Trends

For the Ab-FL complexes, 3PEPS decays are typically observed on at least three different time scales; these have been qualitatively linked to protein flexibility.¹⁵ The fastest decays occur on a ~100 fs timescale and have been assigned to vibrations within local minima of the energy landscape. Dephasing on a ~1 to 10 ps timescale has been assigned to diffusive motions between these minima, and dephasing on even longer timescales to conformational dynamics, some of which is too slow to be time-resolved, thus leading to a non-zero asymptote in $C(\tau)$ at the longest observable τ . When comparing the 3PEPS decays of $V_L^{gl} V_H^{gl}$, $V_L^{4-4-20} V_H^{4-4-20}$, $V_L^{H34R} V_H^{4-4-20}$, and $V_L^{4-4-20} V_H^{4-4-20}$, two major trends are observed: upon affinity maturation, amplitude shifts from long timescale decay components to short timescale components, while the time constants of the picosecond decays tend to become longer.^{12,13} This data is consistent with an increase in rigidity of the Ab combining site. A shift in amplitude to shorter timescales indicates that the protein samples less phase space through conformational dynamics, while increasing time constants of the picosecond decays indicate higher barriers between conformational substates.

The $C(\tau)$ computed for each Ab and the multi-exponential fits are shown in Fig. 4; corresponding τ_j and A_j are shown in Table 1. Since the total length of each simulation was 10

ns, the correlation functions were not computed for time periods exceeding 5 ns. Thus, τ_j values larger than approximately 2 ns denote processes that occur on time scales longer than we can reliably determine from the simulations. Moreover, the distinct reoccurrences in the $C(\tau)$ of $V_L^{gl}V_H^{4-4-20}$ and $V_L^{gl}V_H^{gl}$ for times longer than 100 ps (see Fig. 4) indicate that 10 ns trajectories are not sufficient to capture all of the slower dynamics, since a non-monotonic decay in $C(\tau)$ on this time scale is most likely due to insufficient averaging. Nevertheless, comparison of simulated and experimental $C(\tau)$ shows that the simulations reasonably reproduce the experimentally observed trends (Table 1). For $V_L^{4-4-20}V_H^{4-4-20}$, $C(\tau)$ decays completely within ~ 1 ns, thus the simulations are expected to capture all of the dynamics, and we in fact find an excellent agreement between simulated and experimental $C(\tau)$. For $V_L^{H34R}V_H^{4-4-20}$, $V_L^{gl}V_H^{4-4-20}$, and $V_L^{gl}V_H^{gl}$, the simulated and experimental $C(\tau)$ do not agree as well, certainly due in part to the fact that the simulations apparently do not capture all of the dynamics for these Ab on longer timescales. Nevertheless, the data still reproduces the main experimental observation, *i.e.* that amplitude shifts from longer time scale motions to shorter timescale motions upon maturation. In $V_L^{4-4-20}V_H^{4-4-20}$, 96% of the amplitude of $C(\tau)$ decays on timescales shorter than 100 ps, compared to 54% in $V_L^{H34R}V_H^{4-4-20}$, 84% in $V_L^{gl}V_H^{4-4-20}$, and 77% in $V_L^{gl}V_H^{gl}$. In addition, the overall inhomogeneity of the Ab-FL complexes as measured by the absorption line width¹⁴ is qualitatively reproduced by the covariance s^2 (Table 1). Thus, the MD simulations are sufficiently validated by the experimental results, and should be suitable for deconvolution of $C(\tau)$ into its components arising from motions of individual amino acids.

4.2 Contributors to Dephasing are Distributed throughout the Combining Site

Having validated the MD simulations, we proceed to use the methodology presented in section 2.2 to correlate the experimental signals to specific protein dynamics. For brevity, we will focus on the mature $V_L^{4-4-20}V_H^{4-4-20}$ and germline $V_L^{gl}V_H^{gl}$, *i.e.* inspect the overall changes that result from affinity maturation. To assess how well the experimentally measured signals relate to binding site flexibility, we will address two questions. First, are the experimentally observed dynamics representative of the entire combining site, or do they reflect motions of only a few, strongly coupled residues? Second, does the observed change in the experimental signals upon affinity maturation occur because the dynamics of individual residues change between germline and mature Abs, or because the FL transition dipole samples motions of different residues in the germline and mature Ab due to structural changes brought about by affinity maturation?

To address these questions we will first consider the $C_i^{partial}(\tau)$ computed for ten residues in the combining site that clearly interact with charged atoms in FL (Fig. 5). The associated multiexponential fits are shown in Table 2. Note that the Kabat numbering system is employed for residue designations;⁴⁷ conversion to crystallographic numbering is provided in the Supporting Information.^{14,20}

We find distinct differences in the pico- to nanosecond decays of the $C_i^{partial}(\tau)$ between the examined residues and as a function of Ab maturation. For all but one of the examined residues, the sum of decays in this time regime are of larger amplitude in $V_L^{gl}V_H^{gl}$ than $V_L^{4-4-20}V_H^{4-4-20}$ (Fig. 5), indicating that the phase space sampled by the individual residues decreases upon affinity maturation, in agreement with our results for the global $C(\tau)$. All residues exhibit decays on multiple time scales, and the time constants for the decays in $C_i^{partial}(\tau)$ are rather similar between residues of the same maturation state. For example, each of the ten residues exhibit a

~100 ps decay in the germline Ab, while there is no decay within this time regime for the mature Ab (τ_3 in Table 2).

It is apparent that the time scales present in each of the $C_i^{partial}(\tau)$ are very similar to those observed in $C(\tau)$. Moreover, the general trends seen in the fits extracted from the $C_i^{partial}(\tau)$ are quite similar to those extracted from $C(\tau)$. The relative contribution of high frequency components of the signal increases in each of the residues listed in Table 2 as maturation proceeds. In addition, the individual time constants tend to decrease. Thus, as observed for $C(\tau)$, there is a marked shift towards higher frequency oscillations as a result of maturation. These results indicate that the time scales observed in $C(\tau)$ arise from the contributions of multiple residues and cannot be decomposed into fluctuations from individual residues in the Ab. Thus, all residues participate in the fluctuations that give rise to $C(\tau)$ rather than specific residues giving rise to distinct time scales.

That said, it is insightful to consider the single-residue auto-correlation functions $C_i(\tau)$ and variances s_i^2 shown in Table 3 in order to determine whether the FL transition dipole samples motions of different residues in the germline and mature Ab..

From the ten individual $C_i(\tau)$ and s_i^2 , we constructed the additive correlation functions,

$$C'_{ad}(\tau) = \sum_{i=1...10} s_i^2 C_i(\tau) / \sum_{i=1...10} s_i^2, \text{ according to Eq. (4), and fit them to multi-exponential decays (Table 2).}$$

There are two features of the data shown in Table 3 that should be highlighted. Firstly, as observed for the $C_i^{partial}(\tau)$, the time scales present in each of $C_i(\tau)$ are very similar to those observed in $C(\tau)$. This result indicates that, even when combined in such a way as to enhance the impact of individual residues, the distinct time scales observed in $C(\tau)$ are not separable into contributions from individual residues in the Ab. Thus, all residues participate in the fluctuations that give rise to $C(\tau)$ rather than specific residues giving rise to distinct time scales.

The second interesting feature of the data in Table 3 is the manner in which the electrostatic coupling of each residue to the FL transition dipole changes as a function of maturation. The ten residues in Table 3 are ranked according to s_i^2 in $V_L^{4-4-20} V_H^{4-4-20}$, which is a measure of the electrostatic coupling strength of a given residue to the FL transition dipole. Tyrosine residues Y^H100e and Y^H53 exhibit relatively large values of s_i^2 ; together they provide a large fraction of the initial amplitude to $C'_{ad}(\tau)$ in both the germline Ab and mature Ab. However, the relative magnitudes of s_i^2 are non-negligible for all of the other eight residues examined. Thus, the chromophore exhibits appreciable electrostatic coupling to residues widely distributed throughout the entire combining site (see Fig 7). This indicates it provides a truly representative description of combining site dynamics. These results suggest that $C(\tau)$, and in turn the 3PEPS experiment, provides information about the dynamics of the entire combining site rather than of a few strongly coupled residues.

In addition to changes in the overall binding site dynamics displayed in Table 2, another source of the experimental observation may be that the FL transition dipole samples a differently weighted average of single-residue motions due to the varying s_i^2 values in the germline and mature Abs. This question can be addressed through a comparison of the dynamics of the individual residues in the germline and mature Abs, quantified by the single-residue $C_i(\tau)$ and s_i^2 (Table 3). The coupling strength to the FL transition dipole as measured by s_i^2 changes significantly between the germline and mature Ab for seven of the ten residues, most

prominently for Y^H53 (59.3 cal/mol vs. 28.8 cal/mol) and G^H100f (33.8 cal/mol vs. 19.9 cal/mol). Thus, it is possible that the differential coupling of residues to the FL transition dipole plays some role in modulating the changes in experimental observables in addition to changes in the dynamics of the individual residues upon affinity maturation. With regard to Ab dynamics, the observation of similar time constants at different residues and comparable changes of these time constants upon affinity maturation strongly suggests that fluctuations of the corresponding residues are correlated. We now explore this issue in detail.

4.3. Correlations Due to Ab Motions Increase with Affinity Maturation

The decay of $C(\tau)$ is determined by fluctuations of the Ab combining site. Above we have shown that the timescales of single-residue fluctuations, while similar for residues of the same maturation state (mature or germline), change upon affinity maturation. However, to identify the mechanisms whereby the observed changes are manifest, it is not sufficient to inspect only single-residue motions. Protein flexibility and plasticity are determined by collective motions that arise from the coupling of adjacent residues via hydrogen bonds (H-bonds), hydrophobic packing, steric interactions, or salt bridges. In addition, previous studies indicate that presence of FL increases the levels of correlation between different regions of $V_L^{4-4-20}V_H^{4-4-20}$, demonstrating that correlated Ab fluctuations can be modulated by ligand interactions.^{48,49} Thus, correlations occurring in the Ab-FL complexes need to be considered in order to characterize the differences in flexibility between mature and germline Abs. The patterns of coupled fluctuations occurring in each simulation can be conveniently depicted via a covariance matrix, which identifies correlated motions in each Ab. An increase in the covariance between the motions of single residues indicates increased coupling, and such matrices may be employed to identify the networks of residues that exhibit coupled motions. The MD trajectories were used to compute the covariance matrices for each Ab-FL complex as described in section 3.3; results for $V_L^{4-4-20}V_H^{4-4-20}$ and $V_L^{gl}V_H^{gl}$ are displayed in Fig. 6. It is clear that the fluctuations in the two Abs exhibit different levels of correlation. In general, there are more regions exhibiting prominent correlation in $V_L^{4-4-20}V_H^{4-4-20}$ than in $V_L^{gl}V_H^{gl}$, indicating that the mature Ab-FL complex is more structurally interconnected than the germline Ab-FL complex.

We employ summed squared covariance (SSC: see section 3.3) as a measure of how strongly fluctuations of a single residue correlate to fluctuations of FL or to the various structural subunits of the variable region (*i.e.* CDR loops or framework regions). Fig. 8A shows the differences in the SSC between germline and mature Abs for correlations between single residues and FL. Notably, the SSC between residues in the CDR loops and FL are significantly increased in the mature compared to the germline Ab, while smaller differences are observed for residues in the framework regions. The largest increase was found for residues in V_L CDR1 and V_H CDR3, which have been previously predicted to couple more strongly to FL in the mature Ab based on structural data.¹⁴ Specifically, a H-bond between R^L34 of V_L CDR1 and an enolic oxygen on the xanthenone ring of FL is introduced during affinity maturation by somatic mutation H^L34R, as are increased packing interactions between residues in V_H CDR3 and FL. The increase in correlation between residues of the CDR loops and FL may thus be an expression of the tighter binding of FL in the mature Ab. However, residues displaying increased correlation to FL in the mature $V_L^{4-4-20}V_H^{4-4-20}$ relative to that present in the germline $V_L^{gl}V_H^{gl}$ are found not only in the combining site but are distributed throughout both variable domains, demonstrating that affinity maturation leads to enhanced dynamic coupling throughout the Ab molecule (see Table S4, Supporting Information). This conclusion is further supported by the observed changes in SSC for correlations between residues and structural subunits of the variable region, which indicate varying flexibility of the protein itself. We find an increase in single residue-protein correlations upon affinity maturation for most residues

(Fig. 8B). Increased single residue-protein correlations are not restricted to the CDR loops that directly contact FL, consistent with the widespread distribution of residues displaying increased Ab-FL correlations. For example, large increases in SSC values are observed for residues in framework regions V_L FR3 and V_H FR3, well outside the binding site. Thus, affinity maturation does not act on the dynamics of the binding site alone, but affects the dynamics of the framework regions as well. The observed increases in SSC values stem from increased correlations both within and between the structural subunits. Further analysis shows that the amount of correlation within structural subunits increases significantly for V_H CDR3, V_H FR3, V_L CDR1 and V_L CDR3 with a significant decrease only observed for V_H FR1 (Table S3, Supporting Information). Correlations between structural subunits significantly increase as well, most prominently between several CDR loops, and between CDR loops and framework regions. Thus, it appears that the experimentally observed changes in binding site dynamics are the result of global changes in the dynamics of the variable domains. In this regard, the use of simulations strongly complements the experiments by allowing us to understand this process at the molecular level. The 3PEPS experiments, while conclusively demonstrating that the timescales for decay of $C(\tau)$ change upon affinity maturation, are sensitive only to motions in the vicinity of the chromophore due to the electrostatic weighting. However, the increased correlations observed for the affinity-mature Ab reveals that these motions are strongly affected by changes distal from the binding site. Stronger coupling between different regions of the mature Ab limits the conformational space accessible to the binding site, thus reducing its flexibility. Specifically, coupling within and between selected CDR loops increases upon affinity maturation, while the same CDR loops show increased correlations to FL motions. In addition, correlations are also increased between CDR loops and selected framework regions supporting these loops, suggesting that such couplings contribute to the observed rigidification of the binding site. Such a rigidification of the combining site via long-distance interactions was previously predicted based on the observation that some of the somatic mutations are located >20 Å away from the binding site and appear to introduce interactions that cross-link the β strands of the framework regions.¹⁴ Our results further support this interpretation.

4.4. Specific Interactions are Altered as a Function of Maturation

The analyses above demonstrate altered correlations between different regions of the Ab-FL complexes, but do not reveal how these correlations arise. In order to address this question, we employed the information present in the covariance matrices to identify regions of the Ab-FL complexes which deserve closer scrutiny. To this end, the residues displaying greater correlation to FL in Fig 8A were examined to identify possible mechanisms of increased coupling (Table S4, Supporting Information). This list of residues includes M^H100g, W^H103, Y^H102, V^L46 and R^L34 which are part of a network of structural interactions described in previous studies.²⁰ However, several of these residues were not formerly identified; many are also present in Table 2, suggesting that the increased correlation observed for these residues in $V_L^{4-4-20}V_H^{4-4-20}$ plays a role in generating altered decay behavior of $C(\tau)$ relative to $V_L^{gl}V_H^{gl}$.

As discussed in the previous section, residues throughout V_L CDR1 show a dramatic increase in correlation both to FL and to other residues upon affinity maturation. A residue of particular interest is Y^L32, which forms a H-bond with the phenylcarboxylate of FL. It also participates in a network of H-bonds with N^L28 and N^L30 (Fig. 9). The H-bond between Y^L32 and FL is fairly constant throughout the simulations. Likewise, the presence of a H-bond between Y^L32 and N^L28 shows no distinct pattern as a function of maturation. However, the H-bond between Y^L32 and N^L30 does exhibit a marked trend with maturation, as is apparent from the histograms of the distance between N_δ of N^L30 and the hydroxyl group of Y^L32 shown in Fig. 9. Conformations with a H-bond between Y^L32 and N^L30 are increasingly populated in concert with affinity maturation. This expanded H-bonding population is accompanied by an increase in the covariance value computed for the associated atoms, demonstrating that they become

more strongly coupled. This observation suggests that these interactions in part mediate the increased rigidity which results from maturation, ‘stapling’ together regions of the loop containing V_L CDR1 (also see Fig. 6). Since CDRs play a central role in antigen binding, rigidifying this region of the Ab will contribute significantly to decreasing the flexibility of the entire combining site. This effect may play a role in generating the patterns of correlation shown in Figs. 6 and 8. This hypothesis can be tested by replacing N^{L30} with a residue incapable of forming a H-bond to Y^{L32} ; we anticipate carrying out such an experiment in the future. As the entire variable region is dynamically coupled (see Figs. 6 and 8), we expect that the effect of eliminating this H-bond will propagate throughout the entire combining site. We predict that removing the possibility for H-bonding at this position will increase the flexibility of the combining site and decrease structural and energetic coupling between the combining site and the chromophore. While this site in the Ab is likely not the only position that mediates flexibility of the Ab combining site (*e.g.* see Fig. 8), this example effectively illustrates the specific interactions which may determine flexibility in the Ab and provides a concrete link between the experimental observations via 3PEPS and properties of the Ab at the atomic level. This is especially significant because the observations described above would have been difficult to infer without a detailed molecular description of these systems.

5. Concluding Remarks

The molecular simulation studies described in this work lend support to the assertion that 3PEPS spectroscopy provides a representative measure of protein dynamics in the vicinity of a given chromophore probe. In accordance with the 3PEPS measurements, Ab maturation is observed to be associated with more rapid dephasing time scales in MD simulations. However, the simulation methods possess the unique advantage of providing an atomistic description of the Ab groups involved, allowing for a detailed understanding of the associated molecular processes. The increased rigidity that accompanies Ab maturation is seen to be associated with specific modifications to Ab fluctuations and structural interactions, including H-bonding patterns. Moreover, it is apparent that the decreased flexibility which accompanies maturation is not limited to the combining site, but distributed throughout the entire variable region. In the specific case of affinity maturation of Ab 4-4-20, we find that the binding site of the mature Ab is significantly rigidified compared to the germline Ab, and that the rigidification is brought about by increased coupling within and between CDR loops and framework regions, including residues well outside the binding site. This observation may be functionally relevant. In the naïve $V_L^{gl}V_H^{gl}$, flexibility of the binding site may allow recognition of a more diverse array of ligands, giving the immune system the capacity to respond to a variety of antigens (although with relatively low affinity). Once a *bona fide* antigen has been identified, the mode of binding changes to one that is more appropriately characterized by lock-and-key recognition in $V_L^{4-4-20}V_H^{4-4-20}$. In a highly evolved Ab such as $V_L^{4-4-20}V_H^{4-4-20}$, the ability to discriminate between different ligands is thus made possible by a relatively rigid binding cleft that exhibits a high degree of complementarity to the antigen. This could allow the immune system to neutralize specific antigens very efficiently without cross-reactivity. The approaches outlined above provide a broadly applicable procedure by which the origin of flexibility changes can be identified in other molecular systems. The ability to delineate these phenomena at the molecular level bodes well for the elucidation of general principles that govern the interplay between flexibility and molecular recognition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

I. F. T. was supported by a doctoral fellowship from the La Jolla Interfaces in Science. This work was supported by National Institutes of Health Grant GM37554.

References

1. Notkins AL. Trends Immunol 2004;25:174. [PubMed: 15039043]
2. Patten PA, Gray NS, Yang PL, Marks CB, Wedemayer GJ, Boniface JJ, Stevens RC, Schultz PG. Science 1996;271:1086. [PubMed: 8599084]
3. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC. Science 1997;276:1665. [PubMed: 9180069]
4. Romesberg F, Spiller B, Schultz P, Stevens R. Science 1998;279:1929. [PubMed: 9506942]
5. Li Y, Li H, Yang F, Smith-Gill SJ, Mariuzza RA. Nat Struct Biol 2003;10:482. [PubMed: 12740607]
6. Terzyan S, Ramsland PA, Voss EW, Herron JN, Edmundson AB. J Mol Biol 2004;339:1141. [PubMed: 15178254]
7. Manivel V, Sahoo NC, Salunke DM, Rao KV. Immunity 2000;13:611. [PubMed: 11114374]
8. Sagawa T, Oda M, Ishimura M, Furukawa K, Azuma T. Mol Immunol 2003;39:801. [PubMed: 12617995]
9. Li YL, Lipschultz CA, Mohan S, Smith-Gill SJ. Biochem 2001;40:2011. [PubMed: 11329268]
10. Mohan S, Sinha N, Smith-Gill J. Biophys J 2003;85:3221. [PubMed: 14581222]
11. Acierno JP, Braden BC, Klinke S, Goldbaum FA, Cauerhff A. J Mol Biol 2007;374:130. [PubMed: 17916365]
12. Jimenez R, Salazar G, Baldridge KK, Romesberg FE. Proc Natl Acad Sci U S A 2003;100:92. [PubMed: 12518056]
13. Jimenez R, Salazar G, Yin J, Joo T, Romesberg FE. Proc Natl Acad Sci U S A 2004;101:3803. [PubMed: 15001706]
14. Zimmermann J, Oakman EL, Thorpe IF, Shi X, Abbyad P, Brooks CL III, Boxer SG, Romesberg FE. Proc Natl Acad Sci U S A 2006;103:13722. [PubMed: 16954202]
15. Thielges MC, Zimmermann J, Yu W, Oda M, Romesberg FE. Biochemistry 2008;47:7237. [PubMed: 18549243]
16. Jimenez R, Case DA, Romesberg FE. J Phys Chem B 2002;106:1090.
17. Yan YJ, Mukamel S. J Chem Phys 1988;89:5160.
18. Joo TH, Jia YW, Yu JY, Lang MJ, Fleming GR. J Chem Phys 1996;104:6089.
19. Cho M, Yu JY, Joo T, Nagasawa Y, Passino S, Fleming GR. J Phys Chem 1996;100:11944.
20. Thorpe IF, Brooks CL III. Proc Natl Acad Sci U S A 2007;104:8821. [PubMed: 17488816]
21. Mukamel, S. Principles of Nonlinear Optical Spectroscopy. Oxford University Press; New York: 1995.
22. de Boeij WP, Pshenichnikov MS, Wiersma DA. J Phys Chem 1996;100:11806.
23. Mercer IP, Gould IR, Klug DR. J Phys Chem B 1999;103:7720.
24. Cho BM, Walker RC, Amer H, Mercer I, Klug DR, Gould IR. J Phys Chem B 2005;109:5954. [PubMed: 16851650]
25. Zwier MC, Shorb JM, Krueger BP. J Comput Chem 2007;28:1572. [PubMed: 17342706]
26. Finkelstein IJ, Goj A, McClain BL, Massari AM, Merchant KA, Loring RF, Fayer MD. J Phys Chem B 2005;109:16959. [PubMed: 16853158]
27. Nilsson L, Halle B. Proc Natl Acad Sci U S A 2005;102:13867. [PubMed: 16162674]
28. Halder M, Mukherjee P, Bose S, Hargrove MS, Song XY, Petrich JW. J Chem Phys 2007;127:055101. [PubMed: 17688362]
29. Homoele BJ, Edington MD, Diffey WM, Beck WF. J Phys Chem B 1998;102:3044.
30. Abbyad P, Shi XH, Childs W, McAnaney TB, Cohen BE, Boxer SG. J Phys Chem B 2007;111:8269. [PubMed: 17592867]

31. Merchant KA, Noid WG, Akiyama R, Finkelstein IJ, Goun A, McClain BL, Loring RF, Fayer MD. *J Am Chem Soc* 2003;125:13804. [PubMed: 14599220]
32. Massari AM, Finkelstein IJ, McClain BL, Goj A, Wen X, Bren KL, Loring RF, Fayer MD. *J Am Chem Soc* 2005;127:14279. [PubMed: 16218622]
33. Whitlow M, Howard AJ, Wood JF, Voss EW, Hardman KD. *Protein Eng* 1995;8:749. [PubMed: 8637844]
34. Breneman CM, Wiberg KB. *J Comput Chem* 1990;11:361.
35. Frisch, MJ.; Trucks, GW.; Schlegel, GE.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Zakrzewski, VG.; Montgomery, JA., Jr; Stratmann, RE.; Burant, JC.; Dapprich, S.; Millam, JM.; Daniels, AD.; Kudin, KN.; Strain, MC.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Cioslowski, J.; Ortiz, JV.; Baboul, AG.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, RL.; Fox, DJ.; Keith, T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Andres, JL.; Gonzalez, C.; Head-Gordon, M.; Replogle, ES.; Pople, JA. Gaussian 98 Revision A.10. Gaussian, Inc; Pittsburgh PA: 2001.
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *J Comput Chem* 1983;4:187.
37. MacKerell, AD., Jr; Brooks, B.; Brooks, CL., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and its Parameterization with an Overview of the Program. In: Schleyer, PvR, editor. *The Encyclopedia of Computational Chemistry*. Vol. 1. Wiley; Chichester, U.K.: 1998. p. 271
38. Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. *J Comput Chem* 2009;30:1545. [PubMed: 19444816]
39. Feig M, Karanicolas J, Brooks CL III. *J Mol Graph Modell* 2004;22:377.
40. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J Chem Phys* 1983;79:926.
41. Brooks CL III, Karplus M. *J Chem Phys* 1983;79:6312.
42. Ryckaert JP, Ciccotti G, Berendsen HJC. *J Comput Phys* 1977;23:327.
43. Nose S. *Prog Theor Phys Supp* 1991;103:1.
44. Lee MS, Feig M, Salsbury FR Jr, Brooks CL III. *J Comput Chem* 2003;24:1348. [PubMed: 12827676]
45. Steinbach PJ, Ionescu R, Matthews CR. *Biophys J* 2002;82:2244. [PubMed: 11916879]
46. Brooks BR, JaneXic D, Karplus M. *J Comput Chem* 1995;16:1522.
47. Kabat, EA. *Sequences of Proteins of Immunological Interest*. 5. U.S. Dept. of Health and Human Services Public Health Service National Institutes of Health; Bethesda, MD: 1991.
48. Lim K, Jameson DM, Gentry CA, Herron JN. *Biochemistry* 1995;34:6975. [PubMed: 7766607]
49. Lim K, Herron JN. *Biochemistry* 1995;34:6962. [PubMed: 7766606]

Figure 1.

Simulation system displaying truncated antigen binding fragment (Fab) from 4-4-20 Ab. The light and heavy chains of the Ab are shown in red and blue respectively, except for residues restrained during the simulation which are shown in black. Residues N^L28, N^L30, Y^L32, R^L34 as well as the solvent sphere around the combining site are shown in ball-and-stick representation. The ligand FL is displayed in the center of the combining site.

Figure 2. Schematic representation of the 3PEPS experiment. (A) Experimental geometry. (B) Dephasing and rephasing of chromophores. The discontinued phase traces represent random phase changes due to structural fluctuations. (C) Determination of the peak shift decay from experimental data.

Figure 3. Model of FL displaying charges obtained for the ground (bottom) and excited (top) states overlaid on each atom. Charges on unlabeled atoms in the xanthene ring system are related to charges on labeled atoms via reflection in the z axis.

Figure 4. Multi-exponential fits to $C(\tau)$ for each Ab. In each panel the black solid, red dotted, and green dashed lines denote the computed $C(\tau)$, the multi-exponential fit and the residual, respectively.

Figure 5.

$C_i^{partial}(\tau)$ for the ten residues present in Table 2 in $V_L^{gl}V_H^{gl}$ (gray) and $V_L^{4-4-20}V_H^{4-4-20}$ (black).

Figure 6.

Covariance matrix for FL atoms and C_α atoms in $V_L^{4-4-20} V_H^{4-4-20}$ (upper diagonal) and $V_L^{gl} V_H^{gl}$ (lower diagonal). Only values for residues that were not restrained during the simulations (see Fig. 2) are shown. Consequently, the matrix indices (numbers on the plot) do not correspond to successive indices in the crystal structure. Conversion to Kabat numbering is provided as Supporting Information. The light gray bars labeled F1-4 denote the positions of FR1-4 in V_L and V_H , the black bars labeled C1-3 denote the positions of CDR1-3 in V_L and V_H , the dark gray bars denote the position of FL, and the black dots indicate positions of somatic mutations. The position of residue N^L-30 described in the text is denoted by an arrow.

Figure 7. Residues observed to be correlated to FL. Residues listed in Table 2 are labeled and shown in ball and stick view while FL is shown in thick stick representation. Only part of the combining site is shown for clarity. Note that the residues are distributed throughout the entire combining site.

Figure 8.

Differences in sum squared covariance (SSC) values between mature and germline Abs for correlations (A) between single residues and FL, (B) between single residues and all other residues under consideration (black bars), or all other residues of the same structural subunit (gray bars). The gray boxes indicate the positions of CDR loops, the black dots the positions of somatic mutations.

Figure 9.

Hydrogen bonds that are altered as a function of maturation. Left: Representation of the combining site displaying FL and V_L CDR1 (H-bonds are shown as dashed lines). The viewpoint in this panel corresponds to an approximately 180° rotation about the vertical axis relative to that shown in Fig. 7. Right: Distribution of the distance between N_δ of N^L30 and the hydroxyl group of Y^L32 as a function of maturation. The number in each histogram denotes the covariance value computed for this pair of atoms during each simulation.

Table 1

Multi-Exponential Fits to C (τ)

	s^2 (cal/mol)	A_1	τ_1^a (fs)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	A_4	τ_4 (ps)	A_5	τ_5 (ns)	offset	χ^2
MD simulation													
$V_L^{4-4-20} V_H^{4-4-20}$	332.0	0.68	<100	0.08	1.5	0.20	24	0.04	660			0.00	0.54
$V_L^{H34R} V_H^{4-4-20}$	633.6	0.26	<100	0.05	1.3	0.08	27	0.02	236	0.59	6.4	0.00	0.14
$V_L^{gl} V_H^{4-4-20}$	496.3	0.51	<100	0.08	1.5	0.05	15	0.17	82	0.19	0.2	0.00	0.67
$V_L^{gl} V_H^{gl}$	590.4	0.34	<100	0.05	3.5	0.24	44	0.12	94	0.27	9.6	0.00	0.55
experiment ^b													
$V_L^{4-4-20} V_H^{4-4-20}$	300 ^c	0.87	90	0.10	5.0			0.03	500				
$V_L^{H34R} V_H^{4-4-20}$		0.66	70	0.13	3.2					d	d	0.21 ^d	
$V_L^{gl} V_H^{4-4-20}$	390 ^c	0.51	60	0.14	3.2					0.26	3.7	0.08	
$V_L^{gl} V_H^{gl}$	429 ^c	0.39	50							0.41	3.2	0.20	

^aDue to the time resolution of simulation data (200 fs), the value for τ_1 is not well determined.

^bFrom 13, 14.

^cHalf width full maximum of absorption band.

^dThe offset for $V_L^{H34R} V_H^{4-4-20}$ combines all decays slower than ~300 ps (no nanosecond contributions observed for $V_L^{H34R} V_H^{4-4-20}$).

Table 2

Multi-Exponential Fits to $C_i^{partial}(\tau)$ for $V_L^{4-4-20}V_H^{4-4-20}$ and $V_L^{gl}V_H^{gl}$

Residue	A_1	τ_1^a (fs)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	A_4	τ_4 (ps)	χ^2
Y ^H 100e	0.80	<100	0.17	18			0.03	860	0.52
	0.65	<100	0.14	41	0.18	88	0.03	1600	0.47
Y ^H 53	0.91	<100	0.06	23			0.03	1400	0.87
	0.73	<100	0.18	23	0.04	100	0.04	5400	0.86
S ^L 91	0.97	<100	0.03	8.4					0.79
	0.88	<100			0.12	73			0.80
Y ^L 32	0.98	<100	0.02	16					0.64
	0.87	<100			0.13	90			0.79
G ^H 100f	0.75	<100	0.08	24			0.01	610	0.85
	0.66	<100			0.23	60	0.10	140	0.73
N ^L 30	0.93	<100	0.05	25			0.02	790	0.82
	0.87	<100			0.12	94	0.01	2100	0.63
R ^L 34	0.96	<100	0.02	14			0.02	1100	0.83
H ^L 34 ^b	0.88	<100			0.08	70	0.04	130	0.75
N ^L 28	0.95	<100	0.04	22			0.01	510	0.85

Residue	A ₁	τ_1^a (fs)	A ₂	τ_2 (ps)	A ₃	τ_3 (ps)	A ₄	τ_4 (ps)	χ^2
$V_L^{gl}V_H^{gl}$	0.84	<100			0.15	88	0.01	730	0.81
H ^{L27d}									
$V_L^{4-4-20}V_H^{4-4-20}$	0.95	<100	0.04	27			0.01	750	0.85
$V_L^{gl}V_H^{gl}$	0.87	<100			0.13	81			0.86
S ^{H95}									
$V_L^{4-4-20}V_H^{4-4-20}$	0.95	<100	0.04	27			0.01	1100	0.86
$V_L^{gl}V_H^{gl}$	0.87	<100			0.13	79			0.86

^aDue to the time resolution of simulation data (200 fs), the value for τ_1 is not well determined.

^bSomatic mutation.

Table 3

Multi-Exponential Fits to $C_i(\tau)$ and $C(\tau)_{ad}$ for $V_L^{4-4-20} V_H^{4-4-20}$ and $V_L^{gl} V_H^{gl}$

Residue		S_i^2	A_1	τ_1^a (fs)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	A_4	τ_4 (ps)	χ^2
Y ^{H100e}	$V_L^{4-4-20}V_H^{4-4-20}$	53.9	0.79	<100			0.17	20	0.04	360	0.48
	$V_L^{gl}V_H^{gl}$	61.5	0.60	<100	0.12	2.6	0.11	33	0.16	117	0.71
Y ^{H53}	$V_L^{4-4-20}V_H^{4-4-20}$	28.8	0.88	<100	0.05	6.5	0.03	36	0.05	610	0.83
	$V_L^{gl}V_H^{gl}$	59.3	0.76	<100	0.05	3.1	0.15	34	0.04	2900	0.86
S ^{L91}	$V_L^{4-4-20}V_H^{4-4-20}$	25.4	0.92	<100	0.02	6.9			0.06	460	0.80
	$V_L^{gl}V_H^{gl}$	17.5	0.95	<100	0.01	9.1			0.05	100	0.85
Y ^{L32}	$V_L^{4-4-20}V_H^{4-4-20}$	24.6	0.97	<100	0.02	12			0.01	1100	0.86
	$V_L^{gl}V_H^{gl}$	14.1	0.93	<100	0.01	5.9			0.06	120	0.84
G ^{H100f}	$V_L^{4-4-20}V_H^{4-4-20}$	19.9	0.93	<100	0.05	24			0.01	540	0.86
	$V_L^{gl}V_H^{gl}$	33.8	0.71	<100	0.03	6.9			0.26	94	0.63
N ^{L30}	$V_L^{4-4-20}V_H^{4-4-20}$	19.3	0.95	<100	0.02	6.2	0.01	27	0.02	610	0.86
	$V_L^{gl}V_H^{gl}$	11.0	0.93	<100					0.07	82	0.86
R ^{L34}	$V_L^{4-4-20}V_H^{4-4-20}$	18.7	0.97	<100	0.01	11			0.02	290	0.87
H ^{L34b}	$V_L^{gl}V_H^{gl}$	13.4	0.96	<100					0.04	93	0.84
N ^{L28}	$V_L^{4-4-20}V_H^{4-4-20}$	12.9	0.96	<100	0.03	6.8	0.01	53	0.01	340	0.85

Residue	S_i^2	A_1	τ_1^a (fs)	A_2	τ_2 (ps)	A_3	τ_3 (ps)	A_4	τ_4 (ps)	χ^2
$V_L^{gl}V_H^{gl}$	13.6	0.91	<100			0.01	14	0.08	118	0.86
H ^L 27d										
$V_L^{4-4-20}V_H^{4-4-20}$	11.4	0.97	<100	0.02	5.9	0.01	37	0.01	610	0.87
$V_L^{gl}V_H^{gl}$	11.0	0.94	<100					0.06	110	0.84
S ^H 95										
$V_L^{4-4-20}V_H^{4-4-20}$	10.8	0.97	<100	0.01	10	0.01	56	0.01	620	0.87
$V_L^{gl}V_H^{gl}$	10.5	0.95	<100					0.05	91	0.85
C(τ_{ad}) ^c										
$V_L^{4-4-20}V_H^{4-4-20}$	226	0.90	<100	0.01	2.8	0.06	22	0.03	580	0.58
$V_L^{gl}V_H^{gl}$	246	0.72	<100	0.05	26	0.10	97	0.13	2.2×10^4	0.61

^aDue to the time resolution of simulation data (200 fs), the value for τ_1 is not well determined.

^bSomatic mutation.

^cAdditive correlation function constructed from the ten residues under consideration, $C'_{ad}(\tau) = \sum_{i=1 \dots 10} S_i^2 C_i(\tau) / \sum_{i=1 \dots 10} S_i^2$, S_i^2 is reported in cal/mol.