Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7

Lillian T. Chong*†, Yong Duan†, Lu Wang†, Irina Massova†, and Peter A. Kollman*†‡

*Graduate Group in Biophysics and †Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143-0446

Edited by Peter G. Schultz, the Scripps Research Institute, La Jolla, CA, and approved October 12, 1999 (received for review July 8, 1999)

We investigated the relative free energies of hapten binding to the germ line and mature forms of the 48G7 antibody Fab fragments by applying a continuum model to structures sampled from molecular dynamics simulations in explicit solvent. Reasonable absolute and very good relative free energies were obtained. As a result of nine somatic mutations that do not contact the hapten, the affinity-matured antibody binds the hapten $>10^4$ tighter than the germ line antibody. Energetic analysis reveals that van der Waals interactions and nonpolar contributions to solvation are similar and drive the formations of both the germ line and mature antibody–hapten complexes. Affinity maturation of the 48G7 antibody therefore appears to occur through reorganization of the combining site geometry in a manner that optimizes the balance of gaining favorable electrostatic interactions with the hapten and losing those with solvent during the binding process. As reflected by lower rms fluctuations in the antibody–hapten complex, the mature complex undergoes more restricted fluctuations than the germ line complex. The dramatically increased affinity of the 48G7 antibody over its germ line precursor is thus made possible by electrostatic optimization.

The immune system is capable of producing antibodies that can bind virtually any molecule with high affinity through the combinatorial association of variable, joining, and diversity genes with subsequent affinity maturation (1). Affinity maturation involves repeated cycles of somatic mutations in the variable regions, preferentially stimulating high-affinity antibodies in the resulting diverse pool of antibodies until an optimal affinity is achieved (2, 3). Indeed, nature has solved many of the complex problems associated with molecular recognition by generating and screening a large library of proteins over the course of a few weeks during an immune response. By unraveling the mechanism behind molecular recognition, we will gain a deeper understanding of the principles that lead to high-affinity and selective inhibitors for a target receptor in the area of drug design.

Over the past decade, chemists have tapped the enormous diversity of the immune system to obtain antibody catalysts for chemical reactions of interest by immunizing with transition-state analog haptens. Research involving catalytic antibodies has provided us with structural insights into the process of affinity maturation. Specifically, high-resolution x-ray crystal structures for the germ line and mature Fab fragments of the esterolytic antibody 48G7, by themselves and in complex with the *p*-nitrophenyl phosphonate hapten, have been attained (4, 5). As a result of nine somatic mutations that are distant from the combining site, the mature antibody has a $>10^4$ greater affinity for the hapten with a dissociation constant (K_d) of 10 nM, compared with a K_d of 135 μM for the germ line antibody (6). Conformational changes occur in the variable region of the germ line antibody in response to binding hapten; these changes become preorganized in the combining site of the mature antibody, which undergoes little change in structure on binding hapten. Structural aspects that lead to the dramatic difference in binding affinities between the germ line and mature antibody–hapten complexes are not well understood. The availability of high-resolution structural information on the 48G7 antibody presents the opportunity to perform a detailed energetic analysis of the system.

A practical approach to calculating free energies is one developed by Srinivasan *et al.* in 1998 (7), which involves applying a continuum model to solute configurations sampled as "snapshots" from a molecular dynamics (MD) simulation by using explicit solvent. For each solute configuration, a molecular mechanical energy is determined. Free energies of solvation are estimated by using finite-difference Poisson–Boltzmann (PB) calculations for the electrostatics contribution and a surface-area-dependent term for the nonelectrostatic contribution to solvation. Solute entropic contributions are estimated from a harmonic analysis. This approach has been used to investigate nucleic acid conformational problems, such as the relative stability of the "A" and "B" forms of nucleic acid duplexes (7, 8); it has also been used for a protein–RNA complex (C. Reyes and P.A.K., unpublished observations) and for different mutants of a peptide–protein complex (9). Here the approach of Srinivasan *et al.* (7) is applied to a small molecule–protein complex problem: the relative binding free energies of the 48G7 germ line and mature antibody–hapten complexes.

Methods

Preparation. Coordinates were extracted from the x-ray crystal structures of the 48G7 germ line and mature Fab fragments complexed with the *p*-nitrophenyl phosphonate hapten (1aj7 and 1 gaf for the germ line and mature Fab–hapten complexes, respectively, in the Protein Data Bank) (4, 5). All crystallographic waters were included (151 waters for the mature Fab–hapten complex). Hydrogen atoms were added to the heavy-atom positions of the crystal structures with the *Leap* module of the AMBER 5.0 package (10). His^L189, His^L198, and His^L55 (in the mature Fab–hapten complex only) were protonated at the Δ-nitrogen; His^H35, His^H164, and His^H200 were protonated at the ε-nitrogen [light chain (L) and heavy chain (H)]. Charges for the hapten were obtained by restrained electrostatic potential fitting (11), and the electrostatic potentials were produced by single-point quantum mechanical calculations at the Hartree–Fock level with a 6–31G* basis set. The resulting partial atomic charges of the hapten are shown in supplemental Table 5 (see www.pnas.org). Bond, angle, and dihedral parameters for the hapten not included in the Cornell *et al.* force field (12) are listed in supplemental Table 6.

Simulation. MD simulations were performed by using the Cornell *et al.* force field (12) and the AMBER 5.0 suite of programs (10). Long-range nonbonded interactions were truncated by using a

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

Abbreviations: H, heavy chain of antibody; L, light chain of antibody; MD, molecular dynamics; PB, Poisson–Boltzmann.

‡To whom reprint requests should be addressed at: Department of Pharmaceutical Chemistry, University of California, 513 Parnassus Avenue, San Francisco, CA 94143-0446. E-mail: pak@cgl.ucsf.edu.

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

12-Å residue-based cutoff. The SHAKE algorithm was applied to constrain all bonds to their equilibrium values, thus removing high-frequency vibrations (13). Each system was solvated by placing a spherical cap of TIP3P water molecules (14) with a radius of 25 Å from the geometric center of the hapten. Unfavorable interactions within the structures were relieved with steepest descent followed by conjugate gradient energy minimization until the rms of the elements in the gradient vector was less than 10^{-4} kcal/(mol·Å). Solvent in the systems (including crystallographic waters) was then equilibrated for 25 ps while raising the temperature from 0 K to 300 K. A second equilibration phase of 100 ps at 300 K was applied with the "belly" option in which only those residues within 12 Å of the hapten geometric center were allowed to move. Constant temperature was maintained by the Berendsen coupling algorithm (15) with separate solute–solvent and solvent–solvent coupling. Sampling of reasonable configurations for the given stable states of the antibody–hapten complex structures was conducted by running a 500-ps belly simulation with a 2-fs time step at 300 K.

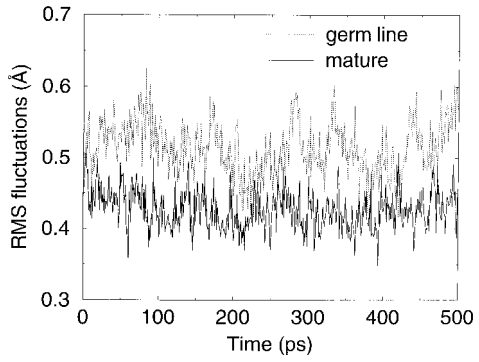


Fig. 1. rms fluctuations of the germ line and mature antibody–hapten complexes. rms fluctuations are defined as rms deviations of the structure at a given time from the average structure of the MD simulation. Only mobile atoms of the simulation, or ''belly'' atoms (see *Methods*), were considered for the rms fluctuations.

Energetic Analysis. The general strategy used here involves calculating energies for "snapshot" configurations taken from the MD trajectories of the antibody–hapten complexes and then averaging the values (7). Unbound antibody and hapten snapshots were also taken from the antibody–hapten complex trajectories. Fifty snapshots taken at 10-ps intervals from each antibody–hapten complex trajectory were used for analysis. For consistency, all waters were removed from each snapshot before energy calculations were performed. Removal of crystallographic waters was not a concern, because none of the waters lie within the combining site.

Total molecular-mechanical energies (E_{gas}) and internal energy (E_{int} where int = bonds, angles, and dihedrals), van der Waals (E_{vdW}), and electrostatic (E_{elec}) components were determined by using the *anal* module of the AMBER 5.0 package (10) with the same force field (12) used in the MD simulations, with the inclusion of solute–solute pairwise interactions. No cutoff was used for the evaluation of nonbonded interactions.

Solute entropic contributions were estimated for the 50th snapshots taken from the trajectories by using the *nmode* module of the AMBER 5.0 package (10), which involves a harmonic approximation to the normal modes and standard (quantum) formulas. Residues lying outside of the belly region in the simulations were removed before minimizing and proceeding with normal mode analysis. All minimizations and normal mode calculations were performed with a distance-dependent dielectric ($\varepsilon = 4r$, where r = interatomic distance in Å) to mimic solvent screening. Steepest descent followed by conjugate gradient minimizations was carried out with a nonbonded cutoff of 12 Å until the rms of the elements in the gradient vector was less than 10^{-4} kcal/(mol·Å). The structures were further minimized with no cutoff for nonbonded interactions by using conjugate gradient and then Newton–Raphson minimizations until the rms of the elements in the gradient vector was less than 10^{-5} kcal/(mol·Å). Normal mode calculations were then carried out with no cutoff for nonbonded interactions.

The electrostatic contribution to the solvation free energy (G_{PB}) was determined with the PB approach (16). This approach involves using a continuum solvent model, which represents the solute as a low dielectric medium ($\varepsilon = 1$) with embedded charges and the solvent as a high dielectric medium ($\varepsilon = 80$) with no salt. Atomic charge values were taken from the Cornell *et al.* force field (12) as consistent with the molecular-mechanical energy calculations. However, atomic radii were taken from the PARSE parameter set (17) instead of the Cornell *et al.* force field (12) because of the small size of hydrogens in the latter. The dielectric boundary is the contact surface between the radii of the solute and the radius (1.4 Å) of a water probe molecule. The DELPHI 2.0 computer program package was used to numerically solve the linearized PB equation on a cubic lattice by using iterative finite-difference methods (16). One thousand iterations were performed for each calculation. The cubic lattice, which had a grid spacing of 0.5 Å, was scaled such that its dimensions were 80% larger than the longest dimension of the solute; the points on the boundary of the grid were set to the sum of Debye–Huckel potentials (18).

The nonpolar contribution to the solvation free energy (G_{np}) was determined by using the solvent-accessible surface area algorithm (SASA) of Sanner (19) in a linear parameterization, where $G_{np} = \gamma$ (SASA) + β and the constants γ and β are 0.00542 kcal/Å^2 and 0.92 kcal/mol, respectively, for use with PARSE atomic radii values (17). The solvent probe radius was set to 1.4 Å.

Computational Mutagenesis. Desired alanine and glycine mutations were made by editing the snapshots from the trajectories before energetic analysis by using an approach developed by Massova and Kollman (9). In this approach, a residue is mutated to an alanine by truncating the residue at Cγ and replacing Cγ with a hydrogen atom at a 1.09-Å distance from Cβ along the Cγ–Cβ bond; likewise, a glycine mutation involves truncating the residue at Cβ and replacing Cβ with a hydrogen atom. The partial charges of the residue were then changed to those appropriate for alanine or glycine by simply changing the topology files. Energetic analysis of the mutants was performed as described above for the wild-type structures.

Results and Discussion

Structure and Dynamics of the Antibody–Hapten Complexes. MD simulations of the 48G7 germ line and mature antibody–hapten complexes were performed with a spherical cap of explicit waters centered at the hapten and a "belly," in which only atoms within a defined distance from the hapten were allowed to move during the simulation. Stable trajectories of the complexes resulted, as demonstrated by steady rms fluctuations (see Fig. 1) and rms deviations from their minimized crystal structures (data not shown). As shown in Fig. 1, the belly atoms in the germ line complex experienced greater rms fluctuations (0.4 to 0.6 Å) than those in the mature complex (0.3 to 0.5 Å). rms deviations of the belly atoms in the complexes from their minimized crystal structures were 1 Å and 0.7 Å, as averaged over the last 100 ps of the simulations for the germ line and mature complexes, respectively. These results suggest that the germ line complex has greater flexibility than the mature complex. The importance of this greater flexibility is underscored by two well-discussed

Table 1. Energetic analysis of 48G7 antibody–hapten complex formations

	Germ line	Mature
$\langle \Delta E_{elec} \rangle$	−304.9 (2.5)	−349.1 (1.5)
$\langle \Delta E_{vdw} \rangle$	−35.9 (0.5)	−35.5 (0.4)
$\langle \Delta E_{mm} \rangle$	−340.8 (2.4)	−384.6 (1.4)
ΔG_{np}	−4.4 (0.0)	−4.2 (0.1)
ΔG_{PB}	317.6 (2.2)	354.0 (1.2)
ΔG_{solv}	313.3 (2.2)	349.6 (1.2)
$\Delta G_{elec,tot}$	12.7 (0.9)	4.8 (0.5)
ΔG_{tot}	−27.5 (0.6)	−35.0 (0.5)
$-T\Delta S$	18.4	19.3
ΔG_{bind}	−9.1	−15.7

Values in parentheses are SEM.

Results are mean energies (in kcal/mol) and SEM from 50 equally spaced "snapshot" configurations of a 500-ps MD simulation $\langle \Delta E_{mm} \rangle$. Each snapshot contains only the solute, which consists of all antibody and hapten atoms, including hydrogens. Definitions of the remaining energetic components are as follows: $\langle \Delta E_{elec} \rangle$ = electrostatic molecular-mechanical energy; $\langle \Delta E_{vdW} \rangle$ = van der Waals molecular-mechanical energy; $\langle \Delta E_{mm} \rangle$ = total molecular-mechanical energy ($\langle \Delta E_{elec} \rangle + \langle \Delta E_{vdW} \rangle + \langle \Delta E_{int} \rangle$); ΔG_{np} = nonpolar contribution to the solvation energy; ΔG_{PB} = electrostatic contribution to the solvation energy calculated by the PB approach; ΔG_{solv} = total solvation energy ($\Delta G_{PB} + \Delta G_{np}$), $\Delta G_{elec,tot}$ = total electrostatic energy ($\langle \Delta E_{elec} \rangle + \Delta G_{PB}$); ΔG_{tot} = total energy without solute entropic contribution ($\langle \Delta E_{mm} \rangle + \Delta G_{solv}$); $-T\Delta S$ = solute entropic contribution, where T = temperature and S = sum of translational, rotational, and vibrational entropies; and ΔG_{bind} = total energy with solute entropic contribution ($\Delta G_{tot} - T\Delta S$). Energetic contributions for the unbound antibody, unbound hapten, and antibody-hapten complex states in both the germ line and mature systems are shown in Tables 5 and 6, respectively, of supplementary material (see www.pnas.org).

Table 2. Energetic analysis of mutant 48G7 antibody–complex formations

	kcal/mol	
	Germ line mutant	Mature mutant
$\langle \Delta E_{elec} \rangle$	−374.0 (2.5)	−370.0 (1.3)
$\langle \Delta E_{vdw} \rangle$	−35.8 (0.5)	−35.4 (0.4)
$\langle E_{mm} \rangle$	−409.8 (2.4)	−405.4 (1.2)
ΔG_{np}	−4.3 (0.0)	−4.4 (0.0)
ΔG_{PB}	386.5 (2.2)	374.6 (1.0)
ΔG_{solv}	382.1 (2.2)	370.3 (1.0)
$\Delta G_{elec,tot}$	12.5 (0.9)	4.6 (0.5)
ΔG_{tot}	−27.7 (0.7)	−35.2 (0.5)

In order to produce mutant germ line and mature antibodies with identical amino acid sequences, the following mutations were made: $Ser^{L}30$, $Asp^{L}55$, $Glu^{H}42$, $Asn^{H}56$, $Asn^{H}76 \rightarrow$ Ala and $Ser^{L}34 \rightarrow$ Gly in the germ line antibody; $Asn^{L}30$, $His^{L}55$, $Lys^{H}42$, $Asp^{H}56$, $Lys^{H}76$, $Thr^{H}78 \rightarrow$ Ala and $Val^{H}55$, $Asp^{H}65 \rightarrow$ Gly in the mature antibody. See Table 1 legend for definitions of the energy components.

notions of affinity maturation (4, 6): (*i*) the germ line antibody must readily adapt to a variety of antigens to efficiently protect the host from disease; and (*ii*) the plasticity of the germ line antibody facilitates its remodeling by somatic mutation into a "mature" conformation that has a tighter fit for the repeatedly introduced antigen.

Energetic Analysis of the Antibody–Hapten Complexes. The results from energetic analysis of 50 equally spaced snapshots taken from each of the two MD simulations are summarized in Table 1. All energetic analysis was done for only a single MD trajectory of the desired antibody–hapten complex with unbound antibody and hapten snapshots taken from snapshots of that trajectory. The binding free energy (ΔG_{bind}) for the complex was estimated by the following expression:

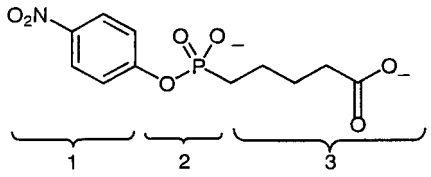
$$\Delta G_{bind} = \langle E_{mm} \rangle + \Delta G_{solv} - T\Delta S,$$

where $\langle E_{mm} \rangle$ is the molecular-mechanical energy, ΔG_{solv} is the solvation energy, and $-T\Delta S$ is the solute entropic contribution. The molecular-mechanical energy consists of internal ($\langle \Delta E_{int} \rangle$), van der Waals ($\langle \Delta E_{vdW} \rangle$), and electrostatic ($\langle \Delta E_{elec} \rangle$) components. Note that, because the structures of each antibody in its unbound and bound states were the same, the internal component of the molecular-mechanical energy has zero contribution to the binding free energy ($\langle \Delta E_{int} \rangle = 0$). The solvation energy consists of an electrostatic contribution (ΔG_{PB}) and a nonpolar contribution (ΔG_{np}). Solute entropic contributions were determined only for the last snapshots of the MD trajectories. The calculated free energies of germ line complex formation ($\Delta G_{bind} = -9.1$ kcal/mol) and of mature complex formation ($\Delta G_{bind} = -15.7$ kcal/mol) agree fairly well with those from experiment (−5.3 kcal/mol and −10.9 kcal/mol, respectively, for the germ line and mature complexes) (6). The favorable formations of these complexes are driven by the van der Waals contributions ($\langle \Delta E_{vdW} \rangle$ values) and the nonpolar contributions to solvation (ΔG_{np} values). These nonelectrostatic components are very similar for both the germ line and mature complex formations ($\langle \Delta E_{vdW} \rangle = -35.9$ kcal/mol and $\Delta G_{np} = -4.4$ kcal/mol for the germ line complex; $\langle \Delta E_{vdW} \rangle = -35.5$ kcal/mol and ΔG_{np} 4.2 kcal/mol for the mature complex). The solute entropic contributions ($-T\Delta S$ values) are similar as well for the germ line and mature complex formations (18.4 kcal/mol and 19.3 kcal/mol, respectively).

It is important to consider the electrostatic component of the molecular-mechanical energy ($\langle \Delta E_{elec} \rangle$) together with the electrostatic contribution to solvation (ΔG_{PB}) when examining the role of electrostatics in the antibody–hapten complex formations. As demonstrated by numerous studies (20–26), electrostatics generally disfavor the docking of ligand and receptor molecules because the unfavorable change in the electrostatics of solvation is mostly, but not fully, compensated by the favorable electrostatics within the resulting ligand–receptor complex. Indeed, the total electrostatic energy contributions ($\Delta G_{elec,tot}$ values) for both the germ line and mature antibody–hapten complex formations are unfavorable, with values of 12.7 kcal/mol and 4.8 kcal/mol, respectively. The mature complex formation is less unfavorable than the germ line complex formation because of a less positive total electrostatic term in which the penalty paid by the electrostatics of solvation is better compensated by favorable electrostatic interactions within the complex. Thus, even though electrostatics destabilize antibody–hapten complex formation, it is the optimized balance of opposing electrostatic contributions that leads to tighter binding in the mature complex.

On the affinity maturation of the 48G7 antibody, changes in either its conformation or amino acid sequence may account for the optimization of electrostatics. In other words, the nine somatic mutations may increase the binding affinity of the antibody for hapten directly through long-range electrostatic interactions or indirectly by inducing a conformational change that leads to improved electrostatic interactions within the immediate binding vicinity. To isolate the effects of conformational change from those of sequence change, we produced identical amino acid sequences for the germ line and mature antibodies while preserving their respective original conformations; this was done by replacing selected residues at the positions of the nine somatic mutations in one or both antibodies to alanines (or glycines, if the residues were glycines in either antibody) with no subsequent minimization. Details of this computational mutagenesis approach are described in *Methods*.

Table 3. Electrostatic interactions between hapten and antibody



	kcal/mol	
	Germ line	Mature
Head (1)	18.4 (0.3)	2.8 (0.3)
Phosphonate (2)	−187.6 (0.9)	−234.4 (0.8)
Tail (3)	−135.8 (2.1)	−117.5 (1.4)
Total	−304.9 (2.5)	−349.1 (1.5)

Results are mean interaction energies (in kcal/mol) and SEM from 50 configurations. Note that the values in the row labeled "totals" correspond to the $\langle \Delta E_{elec} \rangle$ and $\langle \Delta E_{vdW} \rangle$ values shown in Table 1.

We then proceeded to conduct an energetic analysis of the resulting "mutant" antibodies.

Results from our energetic analysis of the mutant germ line and mature antibodies are shown in Table 2. The total free energies (ΔG_{tot} values) for the mutant germ line and mature antibody–hapten complex formations are different, with values of −27.7 kcal/mol and −35.2 kcal/mol, respectively. These values are essentially unchanged from those of the wild-type germ line and mature antibody–hapten complex formations, which are −27.5 kcal/mol and −35.0 kcal/mol, respectively (see Table 1). Conformational differences are therefore key to the affinity maturation of the 48G7 antibody.

This result is not too surprising, because the somatic mutations are distant from the hapten (the closest one being ≈6 Å away). Nonetheless, the similarity of the results for the wild-type and mutant antibody–hapten complexes suggests that long-range electrostatic interactions have little, if any, effect on the formation of the complexes. The difference in the binding affinities of the germ line and mature antibodies for the hapten must therefore lie in the short-range electrostatics, which can be largely influenced by reorganization of the combining site. Electrostatic interactions between the hapten and the residues in the germ line and mature antibodies were thus explored, dividing the hapten into three moieties: a nitroaryl "head" moiety, a phosphonate moiety, and an alkyl "tail" moiety.

As shown in Table 3, the phosphonate moiety of the hapten forms the bulk of the favorable electrostatic interactions between the hapten and the residues in both the germ line and mature antibodies. In fact, the net sum of electrostatic interactions between the antibody and hapten ($\langle \Delta E_{elec} \rangle$) is more favorable for the mature complex ($\langle \Delta E_{elec} \rangle = -349.1$ kcal/mol) than for the germ line complex ($\langle \Delta E_{elec} \rangle = -304.9$ kcal/mol) largely because of a more negative electrostatic term for the phosphonate moiety of the hapten (−234.4 kcal/mol for the mature complex vs. −187.6 kcal/mol for the germ line complex). Note that the electrostatic interactions of the hapten tail are more favorable in the germ line complex (−135.8 kcal/mol) than in the mature complex (−117.5 kcal/mol). However, the differences in the electrostatic interactions of the tail are not as great as those of the phosphonate moiety. The phosphonate moiety is thus a major binding determinant, as suggested earlier by crystallographic studies based on the formation of hydrogen bonds between the phosphonyl oxygens and the combining site residues, $Arg^{L}96$ and $His^{H}35$, in both the germ line and mature complexes; in the mature complex, the phosphonyl oxygens also form an additional hydrogen bond with $Tyr^{H}33$ (4, 5). These crystallographic studies also indicate that the tail of the hapten may be disordered and highly flexible because of its weak electron density relative to that of the phosphonate moiety. This flexibility may introduce error into our calculations. Nonetheless, our results are based on MD simulations that show low rms fluctuations (<1 Å) for the tail region of the hapten in both the germ line and mature complexes.

Table 4. Electrostatic analysis of $Tyr^{H}33$ in 48G7 antibody–hapten complex formations

	kcal/mol	
	Germ line	Mature
$\Delta\Delta G_{PB}$	12.4 (0.8)	8.5 (0.1)
$\langle \Delta\Delta E_{elec} \rangle$	−17.7 (1.2)	−16.5 (0.3)
$\Delta\Delta G_{elec,tot}$	−5.3 (0.5)	−8.0 (0.3)

"Mutant" corresponds to the antibody–hapten complex system with "discharged" $Tyr^{H}33$ (all partial atomic charges in $Tyr^{H}33$ set to zero). See Table 1 legend for definitions of electrostatic energetic terms.

The interaction between the phosphonate moiety of the hapten and $Tyr^{H}33$ is key to the more favorable network of electrostatic interactions between the phosphonate moiety of the hapten and the mature antibody. This interaction, which is worth about −17 kcal/mol, is much more attractive than it is in the germ line complex, where it amounts to about −5 kcal/mol. With such a favorable electrostatic interaction between just the $Tyr^{H}33$ and the phosphonate moiety of the hapten, the electrostatic interactions formed by $Tyr^{H}33$ may actually compensate, or more than compensate, for the penalty paid in the electrostatics of solvation on antibody–hapten complex formation.

These possibilities were addressed by developing a strategy for determining the total electrostatic contribution of an individual residue (i.e., $Tyr^{H}33$), which includes both its favorable electrostatic interactions with the hapten and its unfavorable loss of electrostatic interactions with solvent. This strategy involves creating a "discharged" version of the residue by setting all partial atomic charges within the residue to zero and using the resulting modified topology file for each snapshot taken from the MD simulation of the antibody–hapten complex. One can thus estimate the total electrostatic contribution of the residue rel-

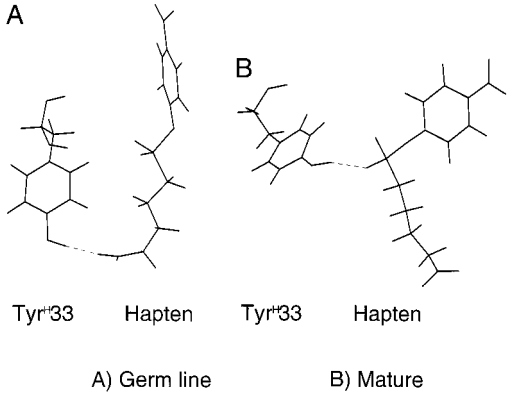


Fig. 2. Orientations of $Tyr^{H}33$ and hapten in the 48G7 germ line and mature antibody–hapten complexes. Orientations are taken from representative structures of the appropriate antibody–hapten complex trajectories. Hydrogen bonds are indicated by dashed lines. (*A*) Germ line antibody–hapten complex. The carboxyl end of the hapten is ideally oriented to form a hydrogen bond with the hydroxyl of $Tyr^{H}33$. (*B*) Mature antibody–hapten complex. Hydrogen bonding is not possible between $Tyr^{H}33$ and the newly positioned carboxyl end of the hapten. However, $Tyr^{H}33$ has shifted toward the phosphonate moiety to enable the formation of a strong hydrogen bond between its hydroxyl and one of the phosphonyl oxygens. Produced with the MIDASPLUS graphics program (26).

ative to its discharged hydrophobic version. Electrostatic contributions were calculated for the resulting "mutants" as described in *Methods* for the wild-type antibody–hapten complexes.

Results for the discharged Tyr^H33 mutants of both the germ line and mature antibodies in complexes with the hapten are shown in Table 4 along with the corresponding values for the wild-type antibody–hapten complexes. As predicted, the total electrostatic contributions of Tyr^H33 to the germ line and mature complex formations are neutral, if not favorable, relative to the discharged version of the residue ($\Delta\Delta G_{elec,tot} = -5.4$ kcal/mol for the germ line complex formation and $\Delta\Delta G_{elec,tot} = -8.1$ kcal/mol for the mature complex formation). Interestingly, the net electrostatic interactions formed by Tyr^H33 in the complex are about the same in the germ line and mature complexes; however, they are distributed differently as a result of the different orientations of the hapten in the two complexes. As shown in Fig. 2, the phosphonate moiety of the hapten is located in similar positions for both complexes, whereas the head and the tail of the hapten are in rather different positions. In the germ line complex, the hydroxyl of Tyr^H33 forms a strong hydrogen bond (worth -13 kcal/mol) with the carboxyl end of the hapten. This hydrogen bond is absent in the mature complex, because the tail of the hapten has oriented itself away from Tyr^H33; however, this missing hydrogen bond is compensated by a strong hydrogen bond (worth -17 kcal/mol) between the hydroxyl of a newly oriented Tyr^H33 and the phosphonate moiety of the hapten. The net $\Delta\Delta E_{elec}$ (Table 4) is actually about 1 kcal/mol more favorable in the germ line antibody, but this is compensated by $\Delta\Delta G_{PB}$, the solvation contribution, which is less favorable than that of the mature complex by about 4 kcal/mol. This more destabilizing electrostatic contribution to solvation is because of greater burial of Tyr^H33 on binding of the germ line antibody to hapten with a solvent-accessible surface area of 45.2 Å^2 in the unbound state and 25.1 Å^2 in the bound state, compared with the values for the mature antibody, which are 22.4 Å^2 and 12.1 Å^2 for the unbound and bound states, respectively.

Approximations and Future Directions. Our approach involving the combination of MD simulations in explicit solvent with the use of continuum solvent models gives reasonable absolute and relative estimates of free energies for the formations of the 48G7 germ line and mature antibody–hapten complexes. These results encourage using this approach as an alternative to Aqvist's linear interaction energy method (27, 28), which also calculates free energies of binding from MD simulations.

A few advantages exist in our approach vs. the linear interaction energy method. One advantage is that our approach requires only one MD simulation, which is for the ligand–receptor complex. The linear interaction energy method requires the generation of two simulations, one for the unbound ligand and one for the ligand–receptor complex. Second, our approach also has the advantage of using the PB treatment of electrostatics, which tends to be more accurate than the linear response approximation for electrostatics used in Aqvist's method. Accurate treatment of electrostatics is particularly important for highly charged systems, such as the 48G7 germ line and mature antibody–hapten complexes, which involve a net charge of -2 on the hapten and net charges of $+3$ and $+5$ for the germ line and mature antibodies, respectively. Thirdly, our approach has the advantage of requiring no empirical parameters that are specific for a given system, whereas Aqvist's method involves an empirical expression for nonpolar effects, which contains parameter values that vary between systems (29).

There are some approximations in our study. First, we have used the snapshots from the MD trajectory of each antibody–hapten complex as representative structures of the unbound antibody and hapten states. Approximating the unbound and bound states of the antibody or of the hapten as the same would be problematic if major conformational changes occur on complex formation. However, only small conformational changes in the binding regions take place in the 48G7 germ line and mature complexes (4). Furthermore, it has been shown for a peptide–protein complex (9) and a protein–RNA complex (C. Reyes and P.A.K., unpublished observations) that, if the true unbound states of the ligand and receptor of interest are used instead of being taken from the complex MD simulations, the resulting absolute binding free energies are of similar magnitudes as when all the free energies are extracted from the single antibody–hapten complex simulation.

Another uncertainty is the harmonic approximation to the solute entropy and its estimation by only a single normal mode calculation. The harmonic approximation involved in normal mode analysis is generally a crude one and, in addition, because of the very shallow potential surface with explicit water, we have removed the waters and done the normal mode analysis with a distance-dependent dielectric ($\varepsilon = 4r$, where r = interatomic distance in Å). We used a distance-dependent dielectric, rather than setting the dielectric equal to one, because minimization with fully charged residues, in the latter case without the waters, is very unrealistic and is sure to move the structure far from the x-ray-determined position. We emphasize that we have used normal mode analysis only to determine approximate estimates of the change in solute entropy on antibody–hapten complex formation.

According to our results, the affinity maturation of the 48G7 antibody involves the optimization of electrostatics; van der Waals interactions appear optimal to begin with in the germ line antibody. These conclusions are consistent with the findings for the Diels–Alderase antibody 39-A11 (30), in which the mature form binds a hydrophobic hapten ≈40-fold tighter than the germ line form. In this situation, the increase in binding affinity is not as dramatic as that for the 48G7 antibody, because the germ line antibody has the ability to form nearly optimal interactions with the hapten. Thus, only two somatic mutations are necessary for the affinity maturation of 39-A11 (30).

As is evident in our energetic analysis of the germ line and mature antibody–hapten complexes, there is a delicate balance between the favorable gas-phase electrostatics term and the unfavorable change in electrostatic contribution to the solvation. In our electrostatic analysis of Tyr^H33, desolvation of the residue on binding of the antibody to the hapten plays a key role in determining the different extents of stabilization provided by the electrostatics of the residue toward complex formation. This type of analysis appears to be useful in isolating the electrostatic effects of individual residues and is likely to uncover more details about the optimization of the 48G7 antibody when applied to other residues in the combining site vicinity.

Conclusions

Using a combination of molecular mechanical energies derived from MD simulations in explicit solvent, solvation free energies derived from a continuum solvent model, and solute entropic contributions derived from normal mode analysis, we have obtained reasonable absolute and very good relative free energies for the 48G7 germ line and mature antibody–hapten complex formations. Energetic analysis reveals that van der Waals interactions and nonpolar contributions to solvation provide the basis for the favorable absolute free energies of the germ line and mature complexes and are optimal even before affinity maturation. The key to the $>10^4$ times tighter binding in the mature complex lies in the electrostatics. By counteracting the favorable electrostatic interactions that form between the antibody and the hapten, the desolvation of the antibody plays an important role in determining the effect of electrostatics, as a whole, on the formation of the antibody–hapten complex.

Alanine/glycine scanning reveals that the interactions of the charged residues involved in somatic mutations are not changed by affinity maturation. Rather, the combining site geometry appears to be reorganized in a manner that optimizes the electrostatics. Increasing the rigidity of the antibody structure further optimizes the binding affinity of the antibody for the hapten. This increased rigidity is apparent in the lower rms fluctuations of the mature complex relative to the germ line complex.

The results of this study have provided us with insight into the process of affinity maturation. First, the immune system must be able to produce high-affinity and specific antibodies rapidly, before the antigen has time to damage the host. In the case of 48G7, this rapid response appears to be achieved through the likely ability of the germ line antibody to adapt and bind to a variety of antigens, because it is more flexible than the mature antibody. Through already optimal hydrophobic interactions but suboptimal electrostatic interactions, the germ line antibody may bind with reasonable affinity to each of these antigens. Greater affinity and specificity are then accomplished through the optimization of electrostatic interactions in the binding site and increased rigidity of the antibody during affinity maturation of the 48G7 antibody.

In the design of an ideal ligand for a given receptor, we must consider the electrostatics and van der Waals interactions as well as the specific arrangement of hydrogen bonds. Optimization, as has been done by the 48G7 antibody, can involve minimizing the desolvation penalty for favorable electrostatic interactions. Such an optimization strategy requires less precision and a more limited set of residues than fine-tuning the precise arrangement of hydrogen bonds and packing of van der Waals interactions. Thus, it is likely that molecular design efforts will benefit from effective strategies for optimizing electrostatics, including desolvation (31, 32), as well as from developing reliable methods for creating ideal hydrogen bonding or van der Waals interactions.

We thank P. Schultz and R. Stevens for providing the crystal structures of the 48G7 antibody system; B. Honig for the DELPHI program; M. Sanner for the solvent-accessible surface area algorithm; and members of our research group, particularly O. Donini, J. Pitera, and C. Reyes, for their helpful suggestions. This work was supported in part by a National Institutes of Health (NIH) Grant GM-29072 to P.A.K. and by a National Science Foundation Fellowship to L.T.C. Graphics were provided by the Computer Graphics Laboratory, University of California, San Francisco (T. Ferrin, prinicipal investigator, NIH P41 Grant RR-01081).

1. Tonegawa, S. (1983) *Nature (London)* **302,** 575–581.
2. Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. (1981) *Cell* **27,** 573–581.
3. Siskind, G. W. & Benacerraf, B. (1969) *Adv. Immunol.* **10,** 1–50.
4. Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. & Stevens, R. C. (1997) *Science* **276,** 1665–1669.
5. Wedemayer, G. J., Wang, L. H., Patten, P. A., Schultz, P. G. & Stevens, R. C. (1997) *J. Mol. Biol.* **268,** 390–400.
6. Patten, P. A., Gray, N. S., Yang, P. L., Marks, C. B., Wedemayer, G. J., Boniface, J. J., Stevens, R. C. & Schultz, P. G. (1996) *Science* **271,** 1086–1091.
7. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. (1998) *J. Am. Chem. Soc.* **120,** 9401–9409.
8. Cheatham, T. E., Srinivasan, J., Case, D. A. & Kollman, P. A. (1998) *J. Biomol. Struct. Dyn.* **16,** 265–280.
9. Massova, I. & Kollman, P. A. (1999) *J. Am. Chem. Soc.* **121,** 8133–8143.
10. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., *et al.* (1997) AMBER 5.0 (University of California, San Francisco).
11. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993) *J. Phys. Chem.* **97,** 10269–10280.
12. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995) *J. Am. Chem. Soc.* **117,** 5179–5197.
13. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977) *J. Comput. Phys.* **23,** 327–341.
14. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983) *J. Comput. Phys.* **79,** 926–935.
15. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984) *J. Comput. Phys.* **81,** 3684–3690.
16. Sharp, K. A. & Honig, B. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 301–332.
17. Sitkoff, D., Sharp, K. A. & Honig, B. (1994) *J. Phys. Chem.* **98,** 1978–1988.
18. Gilson, M. K. & Honig, B. H. (1987) *Nature (London)* **330,** 84–86.
19. Sanner, M. F., Olson, A. J. & Spehner, J. C. (1996) *Biopolymers* **38,** 305–320.
20. Novotny, J. & Sharp, K. (1992) *Prog. Biophys. Mol. Biol.* **58,** 203–224.
21. Novotny, J., Bruccoleri, R. E., Davis, M. & Sharp, K. A. (1997) *J. Mol. Biol.* **268,** 401–411.
22. Misra, V. K., Sharp, K. A., Friedman, R. A. & Honig, B. (1994) *J. Mol. Biol.* **238,** 245–263.
23. Misra, V. K., Hecht, J. L., Sharp, K. A., Friedman, R. A. & Honig, B. (1994) *J. Mol. Biol.* **238,** 264–280.
24. Sharp, K. A. (1996) *Biophys. Chem.* **61,** 37–49.
25. Shen, J. & Wendoloski, J. (1996) *J. Comput. Chem.* **17,** 350–357.
26. Bruccoleri, R. E., Novotny, J. & Davis, M. E. (1997) *J. Comput. Chem.* **18,** 268–276.
27. Aqvist, J., Medina, C. & Samuelsson, J. E. (1994) *Prot. Eng.* **7,** 385–391.
28. Hansson, T., Marelius, J. & Aqvist, J. (1998) *J. Comput. Aided Mol. Des.* **12,** 27–35.
29. Wang, W., Wang, J. & Kollman, P. A. (1999) *Prot. Struct. Funct. Genet.* **34,** 395–402.
30. Romesberg, F. E., Spiller, B., Schultz, P. G. & Stevens, R. C. (1998) *Science* **279,** 1929–1933.
31. Lee, L. P. & Tidor, B. (1997) *J. Chem. Phys.* **106,** 8681–8690.
32. Kangas, E. & Tidor, B. (1998) *J. Chem. Phys.* **109,** 7522–7545.