Ligand Binding in the Catalytic Antibody 17E8. A Free Energy Perturbation Calculation Study

Thomas Fox, Thomas S. Scanlan, and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446

Received December 16, 1996[⊗]

Abstract: We present free energy calculations on the Michaelis complexes of the catalytic antibody 17E8 with two substrates, which differ only in their side chains. Replacing a -CH_2- group with a -S- increases K_M by a factor of about 5−8. This corresponds to a free energy "preference" for the -CH_2- ligand of about 0.9−1.3 kcal/mol. The calculations semiquantitatively reproduce the experimental free energies and show that the preference for a -CH_2- over a -S- is mainly due to the more favorable solvation free energy in the unbound form of the molecule.

I. Introduction

In 1986, the concept of antibody catalysis was reduced to practice with the demonstration that antibodies elicited to transition state analogs showed enzyme-like properties.[1,2] Over the past 10 years, research in the field has demonstrated that this approach is, indeed, general; if a stable mimic of a transition state can be synthesized, it will more often than not lead to antibodies with at least modest levels of catalytic activity.[3−6] More recently, structural biology has taken on a central role in the catalytic antibody field.[7−12] Over the past 2 years, several crystal structures of catalytic antibody Fab fragments complexed with the transition state analog antigen have been solved. These structures provide an opportunity to study specific molecular interactions that are important in catalysis. In particular, an understanding of the specific interactions that relate to transition state complementarity would provide molecular-level insight into the energetics of antibody catalysis and may lead to new transition state analog designs that provide more active catalytic antibodies. Here we report calculational results from a theoretical study aimed at analyzing antibody−transition state analog interactions for an esterolytic antibody.

Recently, Scanlan and co-workers reported the generation and characterization of an unusually active catalytic antibody, 17E8.[13] It is raised against a phosphonate transition state analog of norleucine **1** and mediates the hydrolysis of unactivated norleucine and methionine phenyl esters **2** and **3** (Figure 1). Subsequently, the structure of the antibody Fab fragment complexed with the transistion state analog was solved to 2.5 Å resolution.[9] The overall structure of the 17E8 Fab is very similar to other known antibody structures; the binding loops have adjusted to accommodate the hapten. The phenyl ring of **1** is bound in a hydrophobic pocket. The recognition of the amino acid side chain of norleucine, *n*-butyl, is mediated by the light-chain residues Leu-L89, Gly-L34, and Tyr-L91, which form a hydrophobic binding site.

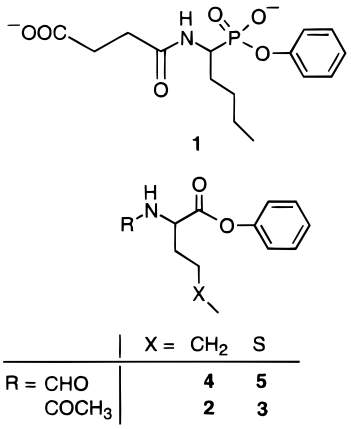


Figure 1. Transition state analog used to raise the catalytic antibody 17E8 (**1**) and the substrates used in the catalysis studies, 2−5.

The antibody shows selectivity for norleucine over methionine as the side chain, where a methylene group is replaced by a sulfur atom. Although this was thought to be a conservative substitution on steric and electronic grounds, the K_M constant for antibody catalysis varied between 5 and 8 times, depending on whether the substituent on the α-amino position bore an acetyl or a formyl group (see Table 1). Assuming that K_M is approximately equal to the dissociation constant of the complex (which is reasonable as the k_{cat} values are very similar), this corresponds to a relative change in free energy of binding of 0.9 or 1.3 kcal/mol, respectively.

To provide a rationale for the experimentally observed results, we decided to perform free energy calculations. An additional goal was to investigate if simulation methodologies are able to reproduce these differences in binding affinity.

* Corresponding author. Tel.: (415) 476-4637. Fax: (415) 476-0688. E-mail: pak@cgl.ucsf.edu.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1997.

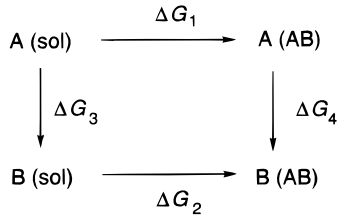
(1) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566−1570.
(2) Pollack, S. J.; Jacobs, J. W.; Schultz, P. G. *Science* **1986**, *234*, 1570−1573.
(3) Lerner, R. A.; Benkovic, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659−667.
(4) Schultz, P. G.; Lerner, R. A. *Acc. Chem. Res.* **1993**, *26*, 391−395.
(5) Hilvert, D. *Curr. Opin. Struct. Biol.* **1994**, *4*, 612−617.
(6) Janda, K. D. *Pure Appl. Chem.* **1994**, *66*, 703−708.
(7) Haynes, M. R.; Stura, E. A.; Hilvert, D.; Wilson, I. A. *Science* **1994**, *263*, 246−252.
(8) Golinelli-Pimpaneau, B.; Gigant, B.; Bizebard, T.; Navaza, J.; Saludjian, P.; Zemel, R.; Tawfik, D. S.; Eshhar, Z.; Green, B. S.; Knossow, M. *Structure* **1994**, *2*, 175−183.
(9) Zhou, G. W.; Guo, J.; Huang, W.; Fletterick, R. J.; Scanlan, T. S. *Science* **1994**, *265*, 1059−1064.
(10) Patten, P. A.; Gray, N. S.; Yang, P. L.; Marks, C. B.; Wedemayer, G. J.; Boniface, J. J.; Stevens, R. C.; Schultz, P. G. *Science* **1996**, *271*, 1086−1091.
(11) Hsieh-Wilson, L. C.; Schultz, P. G.; Steven, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5363−5467.
(12) Charbonnier, J.-B.; Carpenter, E.; Gigant, B.; Golinelli-Pimpaneau, B.; Eshhar, Z.; Green, B. S.; Knossow, M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11721−11725.
(13) Guo, J. C.; Huang, W.; Scanlan, T. S. *J. Am. Chem. Soc.* **1994**, *116*, 6062−6069.

S0002-7863(96)04315-6 CCC: $14.00 © 1997 American Chemical Society

Table 1. Experimental Kinetic Parameters for Substrate Binding to and Ester Hydrolysis in the Antibody 17E8[13]

substrate	K_M (μM)	k_{cat} (min^{-1})
2	215	49.0
3	1033	31.1
4	259	101.4
5	2270	141

Scheme 1



II. Methods

The application of free energy calculations to investigate the relative stability of ligand−protein interactions can be illustrated with the thermodynamic cycle shown in Scheme 1. In Scheme 1, A and B refer to the two different ligands either in solution (sol) or binding to the antibody AB. Experimentally, the free energy difference $\Delta\Delta G$ is measured as $\Delta G_1 - \Delta G_2$. Since free energy is a state function and, therefore, the sum of all energies in this cycle is zero, this is equivalent to $\Delta G_3 - \Delta G_4$. The latter path, however, is preferred computationally, as it does not require the (lengthy) simulation of diffusion of the ligand to the protein or the desolvation—of both ligand and binding sites—upon binding. Assuming that a change in the ligand does not change the binding mode or disrupt the structure of the antibody, both ΔG_3 and ΔG_4 can be calculated with configurational sampling restricted to the local modes at and around the ligand.

In the present study, the thermodynamic integration (TI) method has been employed. Thus, to obtain the free energy difference between two states of a system A and B, we represent the system by a Hamiltonian, $H(\lambda)$. The description of the system, therefore, depends on a coupling constant λ, such that $H(0) = H_A$ and $H(1) = H_B$, where H_A and H_B are the Hamiltonians of states A and B, respectively. Using thermodynamic integration,[14,15] the relative free energy difference between states A and B can be calculated as

$$\Delta G = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda}\, d\lambda = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda$$

where $\langle\ \rangle_\lambda$ denotes the ensemble average at a given λ. In the actual implementation, the integral is evaluated numerically at specified values of λ, which define the "windows" used, and the value of ΔG is obtained by employing the trapezoid rule.

TI allows the formal decomposition of the total free energy into a number of contributions arising from the different terms of the potential energy function, such as bond, angle, dihedral angle, van der Waals, and electrostatic interactions.[16] However, whereas the total free energy is a state function (and the free energy difference between A and B is independent of the choice of the actual integration path over λ), the decomposition and the obtained contributions to the free energy are path-dependent and, therefore, not rigorously defined.[17−19] Nevertheless, even though the free energy components should not be considered as physically exact values, they may be an indication of the relative magnitude of contributions, either from the different terms in the force field or from different regions of the molecular system.[19−21]

Table 2. Additional Force Field Parameters for Substrate Molecules

bonds	K_b (kcal mol^{-1} Å^{-2})	r_0 (Å)
C−OE	450.0	1.364
CT−OE	320.0	1.410

angles	K_θ (kcal mol^{-1} rad^{-2})	θ_0 (Å)
CT−C−OE	70.0	117.0
C−OE−CT	55.0	120.0
O−C−OE	80.0	123.0
OE−CT−H1	50.0	109.5

dihedrals	IDIVF[a]	K_ϕ (kcal mol^{-1} rad^{-1})	phase (deg)	periodicity
X−C−OE−X	1	2.25	180.0	2.0
X−CT−OE−X	3	1.15	0.0	3.0
O−C−OE−CT	1	2.25	180.0	2.0
O−C−OE−CT	1	2.70	180.0	1.0

improper torsions	K_ϕ (kcal mol^{-1} rad^{-1})	phase (deg)	periodicity
OE−O−C−CT	10.5	180.0	2.0

nonbonded parameters	R^* (Å)	ϵ (kcal/mol)
OE	1.6837	0.1700

[a] Factor by which the torsional barrier K_ϕ is divided.

In the present free energy simulation, a -CH_2- group is perturbed into a sulfur atom. The hydrogen atoms of the methylene group are mutated to dummy atoms, and the force field parameters of carbon are changed into those of S using the single topology method.[15]

A. Parameters. The simulations described in this paper were performed with the molecular dynamics simulation package AMBER 4.1,[22] using the all-atom force field by Cornell *et al.*[23]

We decided to investigate the perturbation of the N-acetyl-blocked substrate **2** into its methionine analog **3**. Although the experimentally derived difference in energy of binding between the N-formyl-blocked substrates **4** and **5** is even larger (see Table 1), we did not choose this substrate for the following reasons: first, an acetyl fragment is closer to the groups used to parametrize our force field; second, it is doubtful if our methodology can pick up subtle differences arising from a minor change far away from the mutation site.

To derive the missing dihedral parameters for the ester group, we performed Hartree−Fock calculations with a 6-31G* basis set[24] on methyl acetate and estimated the O−C−OE−C torsion parameters from the *ab initio* rotational barrier. The bond, angle, and improper torsion parameters were chosen in analogy to already existing entries in the force field file. All these additional parameters are given in Table 2.

The partial charges for substrates **2** and **3** were obtained using the standard multiple molecule RESP fit.[25] We calculated the molecular electrostatic potential for the fragments benzoyl acetate (**6**) and methyl- and acetyl-blocked methionine and norleucine (**7** and **8**, see Figure 2) with a 6-31G* basis set[24] and then used these MEPs to derive the atomic charges for the complete substrates. For the simulations, we kept the charges of the phenyl ester and the acetylamine part of the substrates constant (i.e., we used identical charges in both **2** and **3**) and varied only the side chain and C_α charges during the perturbation. The charges are given in Figure 3, together with the AMBER atom types used.

(14) van Gunsteren, W. F.; Weiner, P. *Computer Simulations of Biomolecular Systems;* ESCOM Science Publishers: Leiden, 1989, and references therein.
(15) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(16) Tidor, B.; Karplus, M. *Biochemistry* **1991,** *30*, 3217−3228.
(17) Smith, P. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1994**, *98*, 13735−13740.
(18) Mark, A.; van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167−176.
(19) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25−33.
(20) Boresch, S.; Karplus, M. J. *Mol. Biol.* **1995**, *254*, 801−807.
(21) Sun, Y.-C.; Veenstra, D. L.; Kollman, P. A. *Protein Eng.* **1996**, *9*, 273−281.
(22) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 4.1, University of California, San Francisco, 1995.
(23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J. *Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. Gaussian 92, Revision G.4, Gaussian Inc., Pittsburgh, PA, 1993.
(25) Cieplak, P.; Bayly, C. I.; Cornell, W. D.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357−1377.

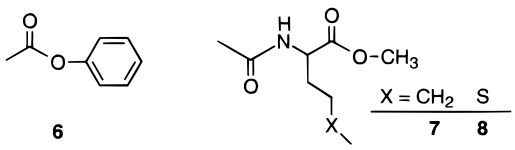


Figure 2. Fragments used to determine the atomic charges of the substrates, benzoyl acetate (**6**) and acetyl- and methyl-blocked norleucine (**7**) and methionine (**8**).

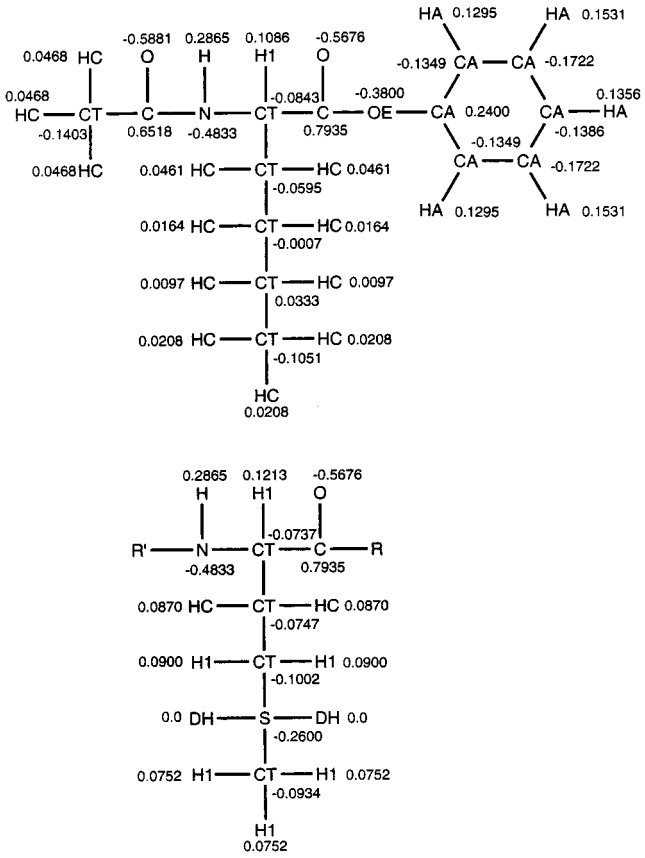


Figure 3. Charges and AMBER atom types of substrates **2** and **3** used in the calculations.

B. Simulations. According to Scheme 1, to determine the relative free binding energies of substrates **2** and **3** to Fab, the free energy changes of mutating **2** into **3** have to be calculated both in aqueous solution and inside the antibody. Additionally, to assess the differences in solvation free energies between the two ligands, we also performed free energy calculations *in vacuo*.

The mutations in vacuum were performed at constant temperature, assigning random velocities that correspond to a temperature of 300 K. We used a time step of 2 fs and SHAKE on all bonds. Every 1000 steps the velocities were reassigned; between reassignments, Berendsen coupling to an external heat bath[26] was used. This is necessary as the usage of SHAKE (which decreases the total energy of the system) in connection with the Berendsen coupling algorithm (to replenish this lost energy) tends to accumulate kinetic energy in selected internal degrees of freedom, such as methyl rotations.[27] For each simulation, we combined a 50 ps equilibration phase with perturbation runs of differing lengths. The 210 ps simulations used 21 windows, each comprising of 1000 steps of equilibration and 4000 steps of data collection, and the 410 (820) ps simulations utilized 41 windows with 1000 + 4000 (2000 + 8000) steps each. Attempts with a larger number of windows but fewer steps per window led to poorer convergence behavior and a significantly increased hysteresis between the forward and the reverse runs (data not shown).

For the solution calculations, substrate **2** was placed in a periodic box of 575 TIP3P water molecules.[28] We used a time step of 1 fs, and the temperature was held at 300 K by means of the Berendsen coupling algorithm,[26] using a coupling constant of 0.2 ps^{-1} and separate coupling of the solvent and the solute. With an inverse compressibility of 44.6 $\times 10^{-6}$ bar^{-1} [29] and a coupling constant of 0.4 ps^{-1}, the pressure was kept constant at 1 atm. A residue-based, switched cutoff for noncovalent interactions of 8 Å was employed. We used SHAKE on all bonds and updated the pair list every 10 steps.

After initial minimization and heating, we performed 50 ps of equilibration and 205 ps of perturbation in each direction. Each perturbation consisted of 41 windows with 1000 steps of equilibration and 4000 steps of data collection each. To test the reliability of the results, we additionally performed perturbations with 41 windows and 2000 + 8000 steps per window, resulting in a total simulation time of 410 ps.

The starting point for the simulations in the catalytic antibody was the crystal structure of the Fab domain with the phosphonate ester **1** as transition state analog.[9] In the coordinate set provided, we replaced the PO_2 group of the analog with CO and shortened the carboxylic acid tether to arrive at substrate **2**. To neutralize the overall charge of the antibody−substrate complex, we used the AMBER4.1 utility program CION[30] to place 14 Cl^- and 9 Na^+ counterions around the antibody. Then a 22 Å water cap was built around the C_α atom of the substrate. Only residues that were within 12 Å of the substrate were allowed to move. Besides increasing computational efficiency, this also prevents water molecules from evaporating from the solvation sphere. To increase computational efficiency further, we also removed the lower half of the Fab domain (residues Ser-124 to Leu-221 in the heavy chain and Lys-107 to Cys-214 in the light chain) in our simulations.

After heating the solution to 300 K, we allowed it to equilibrate for 50 ps and then started the perturbation run. The perturbations in both directions were 205 ps long, with 41 windows of each 1000 steps of equilibration and 4000 steps data collection. All other simulation parameters were the same as in the solvent simulations described above.

As shown in the work on free energy calculations,[31] when the mutation involves changes in the bond lengths, it is necessary to evaluate the free energy of these changes by constraining the bond lengths and to calculate the PMF correction to obtain reliable free energies.[31,32] In essence, the PMF contribution corrects for the free energy change due to lengthening or shortening bonds. This contribution also contains electrostatic and van der Waals terms that change with the bond length. Therefore, the PMF contribution reported in this paper always is, to some extent, coupled with the nonbonded contributions, and some part that is reported as PMF should really be counted as a part of the electrostatic or van der Waals contribution (if we were able to extract those).

The free energies reported in the remainder of this paper always refer to the "forward" direction, i.e., for the mutation of CH_2 into S, regardless of the actual direction of the perturbation run. As a rough error estimate, we use half the difference between the forward and the reverse free energy changes (hysteresis), and the error of $\Delta\Delta G$ is calculated as the square root of the sum of the squares of the individual errors in the free energies.

For each simulated trajectory, we present two sets of free energy data: one that does not include energy contributions from within the perturbed group (which is the whole substrate) and also one set where the intraperturbed group nonbond, internal, and 1-4 energies have been accumulated in the overall free energy change. The PMF correction, however, will be added in both data sets.

Ideally, both the calculations with and those without these contributions should lead to the same results for the free energy change, as one could expect the internal contribution to (almost completely) cancel in

(26) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(27) Simmerling, C., private communication.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(29) Weast, R. C., Ed. *CRC Handbook Chemistry and Physics;* CRC Press: Boca Raton, FL, 1988.
(30) Utility program CION, part of the AMBER 4.1 distribution (ref 22).
(31) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4545.
(32) Pearlman, D. A. *J. Chem. Phys.* **1993**, *98*, 8946−8957.

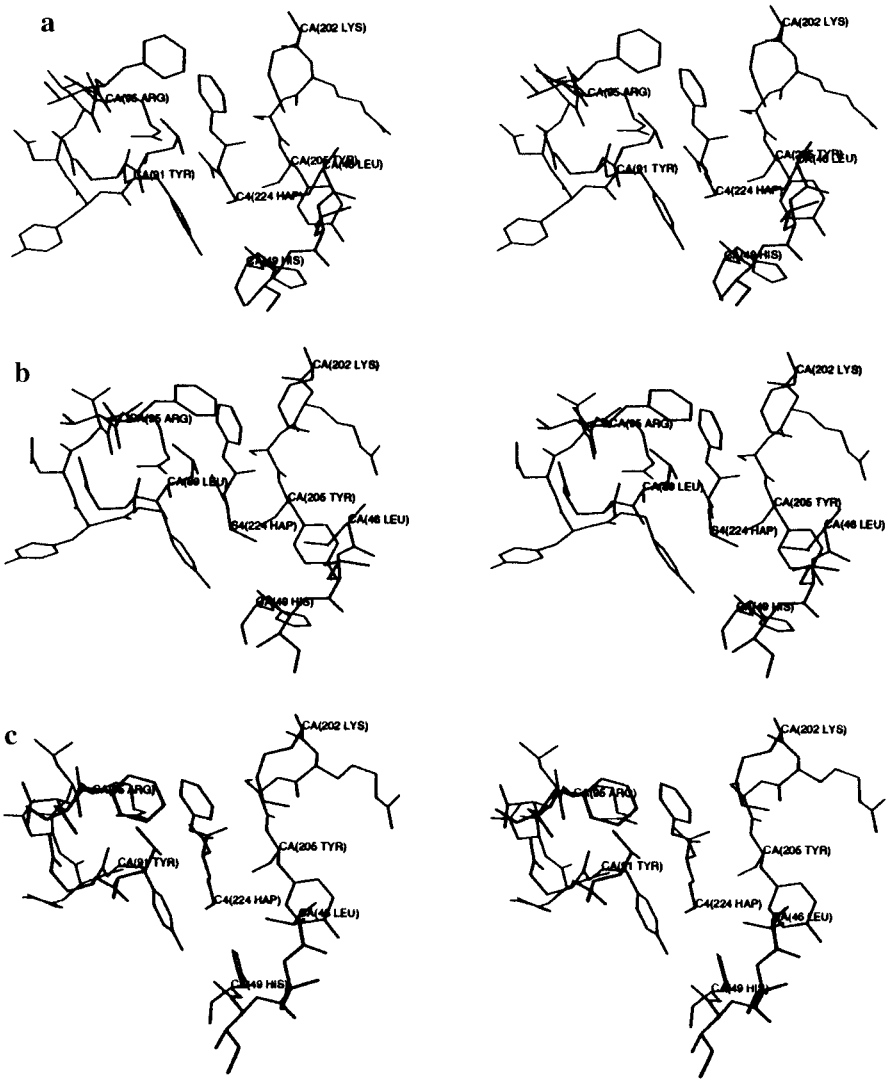


Figure 4. Stereoviews of important Fab residues (a) after initial equilibration (showing **2**); (b) after the reverse perturbation (showing **3**); (c) after the forward perturbation (showing **2**). Residue numbers shown are absolute numbers; residues 1−105 refer to amino acids from the light chain, and numbers greater than 105 refer to amino acids from the heavy chain (e.g., Lys-202 corresponds to Lys-H97).

Table 3. RMSD from Crystal Structure for the Moving Belly of the Fab−Substrate Complex at Various Points of the MD Trajectory

	time (ps)	RMSD (Å) heavy atoms	RMSD (Å) backbone atoms
start	0	0.0	0.0
after minimization and heating	30	0.71	0.53
after reverse equilibration	80	0.75	0.57
after reverse perturbation	280	1.11	0.71
after forward equilibration	330	0.94	0.61
after forward perturbation	530	1.08	0.71

the different legs of the simulation. Therefore, the comparison of these two calculations should be an additional means to assess the overall error of our free energy results.

III. Results and Discussion

Our first check as to the reliability of the simulation was to see if the structure of the antibody−substrate complex stayed close to the crystal structure under the simulation conditions. To do so, we calculated the root-mean-square deviation (RMSD) from the initial structure for all heavy and backbone atoms of the complex that are moving during the simulation. As the results in Table 3 show, the structure stays very close to the initial structure over the whole course of the simulation. After the first equilibration, the heavy atom RMSD amounts to 0.75 Å; this value increases slightly during the reverse simulation to 1.1 Å and then stays roughly constant at about 1.0 Å. This result indicates that the calculated free energy differences are due to the mutations in the ligand and not to some unwanted large conformational changes in the antibody. In Figure 4 we show a stereoview of the Fab residues in the vicinity of the bound substrate after the reverse and after the forward perturbations, which also confirms that, generally, only small changes in the geometry take place.

The calculated free energies without contributions from within the perturbed group of the mutation of substrate **2** into substrate **3** are listed in Table 4.

In vacuum, this perturbation is favorable by about 1.7 ± 0.1 kcal/mol; as there are no interactions with other molecules, this gain is exclusively due to the PMF contribution from lengthening the C−C to S−C bonds. Although the hysteresis between forward and backward runs is relatively small, with longer simulation times the free energy difference in the forward runs is increasing, whereas ΔG for the reverse simulation stays constant at about −1.6 kcal/mol. This finding reflects the difficulties in achieving well-converged vacuum simulations noted elsewhere,[33] probably due to the lack of efficient

(33) Simmerling, C., work in progress.

Table 4. Calculated Free Energies (Sign Corresponds to the Forward Direction, in kcal/mol) for the Perturbation of Substrate **2** into Substrate **3** in Vacuum, in Solution, and in the Fab[a]

run	length (ps)	tot	elec	vdW	pmf
		In Vacuum			
forward	210	−1.45			−1.45
reverse	210	−1.57			−1.57
forward	410	−1.55			−1.55
reverse	410	−1.58			−1.58
forward	820	−1.71			−1.71
reverse	820	−1.61			−1.61
average	820	−1.66 ± 0.05			
		In Solvent			
forward	205	−3.38	−1.59	−0.94	−0.85
reverse	205	−3.31	−1.38	−0.97	−0.96
average	205	−3.35 ± 0.03			
forward	410	−2.78	−1.28	−0.85	−0.65
reverse	410	−3.22	−1.46	−0.87	−0.89
average	410	−3.00 ± 0.22			
		In Antibody			
forward	205	−1.90	−0.63	−1.57	0.29
reverse	205	−0.76	0.23	−1.39	0.39
average	205	−1.33 ± 0.57			
ΔΔ*G* (sol/vac)		−1.34 ± 0.23			
ΔΔ*G* (AB/vac)		0.33 ± 0.57			
ΔΔ*G* (AB/sol)		1.67 ± 0.61			

[a] No intra-perturbed group contributions have been added to the values shown (the entire substrate is considered the perturbed group). Shown are the total change in free energy (tot), as well as the contributions from the electrostatic and the van der Waals energies (elec and vdW). Additionally, we give the value for the bond PMF contribution (pmf).

mechanisms to distribute energy between the different internal degrees of freedom. However, as this energy drift is rather small, we decided against even longer simulations.

The perturbations in water show that the substrate with the methionine side chain is 3.0 ± 0.2 kcal/mol more stable than **2**. About half of this energy gain is due to more favorable electrostatic interactions (Δ*G*(elec) = 1.4 ± 0.1 kcal/mol), whereas van der Waals and PMF contributions account for the other half.

Combining these results with the vacuum numbers yields a differential solvation free energy of −1.34 ± 0.23 kcal/mol for the perturbation of **2** into **3**. This result is intuitively reasonable, as one would expect the sulfur atom to be more capable of interacting favorably with the aqueous solvent than does the CH_2 group. This change in free energy, however, is significantly smaller than the relative solvation free energies of the propane/dimethyl sulfide pair, which can been estimated to be about 3.5 kcal/mol.[34] Obviously, the large hydrophobic portion of the substrate, especially the phenyl ring, either leads to a significant stabilization of the hydrophobic norleucine side chain over the sulfur analog or prevents the S atom from interacting effectively with the solvent. This is also very consistent with results for the mutation of methylserine into norvaline, where the calculations yielded a free energy change that is much smaller than the difference between propane and dimethyl ether.[35]

In contrast to the calculations in vacuum and in solution, the perturbation in the antibody fragment shows a large hysteresis between the forward and the backward runs, which is mainly due to the electrostatic contribution that varies from −0.63 kcal/mol in the forward to +0.23 kcal/mol for the reverse direction. Of course, this points to the problem of a relatively short simulation time of 205 ps, which is probably not long enough for a thorough sampling of important parts of the phase space. However, due to the demanding nature of the problem, we were not able to run significantly longer simulations, which should alleviate this sampling problem.

The difference in binding free energies between **2** and **3** is calculated to be 1.67 ± 0.61 kcal/mol, which—given the uncertainties in our calculations and the large hysteresis—is in good agreement with the experimental value of 0.9−1.3 kcal/mol. The main contribution to the overall free energy change stems from van der Waals interactions, whereas the electrostatic part—especially in the light of its large hysteresis—is essentially zero.

In Table 5, we present results obtained from the same MD trajectory used to generate the numbers in Table 4; however, this time the intraperturbed group contributions are included in the data shown.

The vacuum calculation gives, for the 810 ps simulation, a free energy change of −2.7 ± 0.04 kcal/mol for perturbing substrate **2** into **3**. This is practically the same result as that obtained with the shorter simulation of 410 ps. Moreover, whereas the shorter simulations show a considerable hysteresis for the individual contributions, these variations are small in our longest simulation.

In aqueous solution, both the shorter and the longer simulations yield a free energy change of 3.7 kcal/mol with a hysteresis of 0.1 and 0.2 kcal/mol, respectively. As in the vacuum runs, the two electrostatic contributions are the dominant ones; however, they have opposite signs and, therefore, cancel each other to a large extent. In contrast, the contributions from the van der Waals interactions are much smaller, but they are both in the same direction. As a result, both electrostatic and van der Waals contributions contribute roughly equal amounts to the overall free energy change.

The relative solvation free energy, as the difference between the free energy change in solution and in vacuum, is −1.0 ± 0.2 kcal/mol. This is 0.35 kcal/mol smaller than the result obtained without the inclusion of the intraperturbed group contributions. However, considering the error bars in these calculations, the two results are not significantly different.

The perturbation in the Fab fragment again shows a large hysteresis between the forward and the backward runs, due to a large difference in the electrostatic contribution. The free energy change of 2.2 ± 0.6 kcal/mol leads, together with the result form the perturbation in water, to a binding free energy difference of 1.5 ± 0.6 kcal/mol between **2** and **3**. This is a little lower than the result without the contributions from the perturbed group but closer to the experimental numbers.

The results from Tables 4 and 5 suggest that the main reason for the differences in binding energies of **2** and **3** to the antibody are differences in solvation of the free ligand in aqueous solution. Independent of whether we include the intraperturbed group energies or not, the free energy changes in vacuum and in the antibody are of similar magnitude, whereas it is considerably larger in solution, such that it significantly favors the solvation of **3** with the methionine side chain.

To further understand the factors influencing the selectivity of the antibody binding to **2** and **3**, we analyzed the contributions to the total, van der Waals, and electrostatic free energy arising from different amino acids in 17E8. Again, the caveat made above holds: these contributions from individual residues are—unlike the total free energy—not state functions and, therefore, can give only a crude estimate of their actual

(34) The experimental difference in solvation free energies between $CH_3CH_2CH_3$ and CH_3OCH_3 is −3.9 kcal/mol. Calculation of the mutation of CH_3OCH_3 into CH_3SCH_3 yields a free energy change of +0.4 kcal/mol. Thus, one can estimate the relative solvation free energy between propane and dimethyl sulfide to be about −3.5 kcal/mol.

(35) Veenstra, D.; Kollman, P. A. *Protein Eng.*, submitted for publication.

Table 5. Calculated Free Energies (Sign Corresponds to Forward Direction, in kcal/mol) for the Perturbation **2** → **3** in Vacuum, in Solution, and in the Antibody Fragment[a]

run	length (ps)	tots	elec	vdW	14vdW	14el	badh	pmf
			In Vacuum					
forward	210	−2.93	1.79	−0.11	−0.60	−3.29	0.73	−1.45
reverse	210	−2.90	3.63	−0.82	−0.50	−4.44	0.81	−1.57
forward	410	−2.52	4.88	−0.67	−0.38	−5.58	0.77	−1.55
reverse	410	−2.74	4.09	−0.35	−0.46	−5.16	0.73	−1.59
forward	820	−2.70	3.99	−0.60	−0.41	−4.80	0.83	−1.71
reverse	820	−2.63	3.82	−0.60	−0.50	−4.60	0.86	−1.60
average	820	−2.67 ± 0.04						
			In Solution					
forward	205	−3.61	4.82	−1.13	−0.45	−6.66	0.66	−0.85
reverse	205	−3.87	3.28	−1.26	−0.56	−5.12	0.75	−0.96
average	205	−3.74 ± 0.13						
forward	410	−3.44	4.00	−1.16	−0.52	−5.83	0.73	−0.65
reverse	410	−3.87	2.69	−1.22	−0.53	−4.65	0.73	−0.89
average	410	−3.66 ± 0.21						
			In Antibody					
forward	205	−2.69	6.28	−2.05	−0.57	−7.30	0.65	0.29
reverse	205	−1.64	7.20	−1.91	−0.53	−7.34	0.55	0.39
average	205	−2.17 ± 0.57						
$\Delta\Delta G$(sol/vac)		−0.99 ± 0.21						
$\Delta\Delta G$(AB/vac)		0.50 ± 0.57						
$\Delta\Delta G$(AB/sol)		1.49 ± 0.61						

[a] The intra-perturbed group contributions have been accumulated in the free energy numbers. Shown are the total change in free energy (tot), the van der Waals (vdW), the 1-4 van der Waals and electrostatic energies, (14vdW and 14el), and the internal contributions from bond, angle, and dihedral terms (badh), as well as the value for the bond PMF contribution (pmf).

Table 6. Components of the Free Energy on a Per Residue Basis (in kcal/mol)[a]

	reverse			forward		
residue	total	vdW	elec	total	vdW	elec
Ile-L48	−0.17	−0.01	−0.16	−0.23	−0.03	−0.20
His-L49	−0.48	−0.08	−0.40	−0.76	−0.11	−0.65
Tyr-L91	0.12	−0.18	0.29	0.00	−0.24	0.24
Arg-L96	0.58	0.00	0.58	0.53	0.00	0.53
Lys-H97	0.35	0.00	0.35	0.14	0.00	0.14
Tyr-H100	−0.22	−0.02	−0.20	−0.24	−0.02	−0.21

[a] Shown are all residues that contribute more than 0.20 kcal/mol to the electrostatic, van der Waals, or total free energy change.

importance. However, despite the fact that the component analysis can be path dependent,[17,18] one of the important uses of component analysis[19−21] is to suggest further experiments. The calculated electrostatic or van der Waals free energies can be used to suggest site-specific mutants that would tend to either reduce the binding preference for **2** over **3** or enhance it.

Only a few residues have a van der Waals contribution above 0.1 kcal/mol; in particular, these are Leu-L46, His-L49, Leu-L89, and Tyr-L91. They all are part of the hydrophobic pocket enclosing the side chain of the substrate, and all favor methionine over norleucine.

In contrast to this, more than a dozen Fab residues contribute significantly to the electrostatic part of the free energy change. However, as some contributions favor the CH_2 group and others the sulfur atom, the overall result leads to almost complete cancellation and only a small electrostatic contribution to the free energy change.

As Table 6 shows, residues His-L49, Tyr-L91, Arg-L96, and Lys-H97 are especially important for the electrostatic part of the free energy change. Whereas His-L49 and Tyr-L91 are part of the side-chain binding pocket (see Figure 4), and a large contribution is not unreasonable, Arg-L96 and Lys-H97 are close to the phenyl ring, and ester group and large contributions from these residues are surprising. With the exception of residues HisL49, Lys-H97, and Asp-H107 and the substrate itself, the results for forward and reverse runs are very similar; the main difference in these cases stems from the electrostatic contribution.

Overall, whereas His-L49 and—to a lesser extent—Ile-L48 and Tyr-H100 strongly favor methionine over norleucine as a side chain, Arg-L96 and Lys-H97 favor **2** over **3**. The results of these component analyses can be tested by site-specific mutation experiments, where these residues are changed to Ala.

Nonetheless, we must emphasize that free energy components, unlike the total free energy, can be sensitive to the pathway chosen in free energy calculations. This fact has led van Gunsteren and co-workers[17,18] to suggest that breaking the free energy into components is inappropriate or useless. A different point of view is found in papers by Karplus and co-workers[19,20] and Sun et al.,[21] who have suggested that free energy component analysis is still beneficial because, if properly interpreted, it can give physical insight and be useful (i.e., suggest experiments). Also, Brady et al.[36] have analyzed which components are correlated or uncorrelated. Thus, there has been considerable effort in the literature to analyze whether it is worthwhile to break the total free energy into components, and our opinion, as also noted in ref 21, is that it is usually useful to do so.

We now turn to a critical analysis of our methodology. The methods used in the free energy calculations presented here have been applied to a wide variety of molecular systems.[15] Nonetheless, we should also note the major difficulties and weaknesses with such calculations, many of which have been pointed out by the reviewers of this paper and are also being discussed in the current literature.

First, the use of a simple residue-based cutoff of 8 Å is clearly a severe approximation. Obviously, a more rigorous approach would be to employ a larger cutoff, reaction field, Ewald, or fast multipole method to include long-range electrostatic effects, and this is certainly warranted for large changes in polarity or charge during the perturbation. Nonetheless, for relatively

(36) Brady, G. P.; Szabo, A.; Sharp, K. A. *J. Mol. Biol.* **1996**, *263*, 123−125.

nonpolar perturbations such as those considered here, a simple cutoff appears reasonable.

Second, our force field is a state-of-the-art effective two-body one, although more sophisticated ones that include polarization effects explicitly could be employed. However, these are likely to be more important for highly polar/charged systems than those considered here and have the disadvantage that they require significantly more computational resources.

Third, we have based our statistical errors on multiple simulations, but more rigorous methods to estimate these errors exist.[37] Nonetheless, in our experience, these more rigorous methods tend to underestimate the real uncertainties.

Fourth, there are other simulation details which could affect the calculated results, such as the use of Berendsen et al.[26] rather than Nose−Hoover[38,39] temperature coupling, or the use of a trapezoidal rather than Simpson's rule in the numerical integration of the free energies. However, it is unlikely that employing those methods will significantly change the free energies.

In fact, as noted in ref 15, the single largest uncertainty in the calculated results, which overshadows the issues noted above, is sampling. By using a highly constrained protein site, in which only residues within 12 Å of the ligand are allowed to move, we are in some sense "forcing" the protein−ligand complex to stay near the X-ray structure. Allowing more flexibility is not necessarily more rigorous, since, with force field/long-range cutoff inaccuracies, it is likely this will lead to further drift from the correct structure and, therefore, sampling in the "wrong" part of phase space. Even in solution, we cannot claim full sampling of all the conformations of **2** and **3**, just a sampling of a subset that is representative of the bound conformation. One reviewer has noted that one could take advantage of more cancellation of errors if one used similar boundary conditions in the ligand-free and ligand-bound trajectories. This may be true, but, in our opinion, one would like to calculate the solvation free energy as accurately as possible, since this is also an experimentally measurable quantity, so we prefer to use periodic boundary conditions in the ligand-free simulations. Using periodic boundary conditions for the ligand-bound trajectory is the most rigorous way to study it, but, given the large size of the antibody, a constraint model of the system is currently a necessary compromise. As noted above, the only absolute necessity is to stay near the crystal structure in the ligand binding site and to allow a modicum of flexibility in the model. The success of such an approach in the example presented here and in others in ref 15 suggests it is not an unreasonable one.

Despite all the uncertainties pointed out above, it is encouraging that our calculated relative free energy of ligand binding based on a number of independent simulations is in good agreement with experiment, independent of the inclusion of intragroup free energy contributions. Of course, this *per se* does not mean our model is "validated". However, when taken together with many such agreements,[15] it should at least be considered a reasonable one, which can be used to rationalize and interpret experimental results, or to predict interesting effects that would induce subsequent experimental efforts.

IV. Conclusions

We performed free energy perturbation calculations on the binding of two ligands, one with a norleucine and the other with a methionine side chain, to the catalytic antibody 17E8. Experimentally, the binding of norleucine is favored by 0.9−1.3 kcal/mol. Our calculations yield—depending on the treatment of the intraperturbed group contributions—values of 1.67 ± 0.61 and 1.49 ± 0.61 kcal/mol. Considering the large error bars in the calculations, this is in very good agreement with the experimental results. Comparison of the free energy changes for ligand **2** to **3** in vacuum, in aqueous solution, and in the antibody fragment suggests that the preferred binding of the ligand with the norleucine side chain is mainly due to the more favorable interaction of the methionine analog with the solvent. A free energy component analysis showed that, in the solvent, both electrostatic and van der Waals terms contribute almost equal parts to the total free energy. In the antibody, however, the van der Waals part clearly dominates over the electrostatic contributions. Although several amino acids of the Fab fragment give large free energy contributions, these terms are of different sign and, therefore, cancel each other to a large extent. The van der Waals contributions, however, all favor the methionine analog inside the antibody.

Acknowledgment. T.F. is grateful for a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. P.A.K. was supported by the NSF (Grant CHE 94-17458) and by the NIH (Grant GM 29072). T.S.S. is supported by the NIH (GM 50672).

JA964315J

(37) Hodel, A.; Simonson, T.; Fox, R.; Brunger, A. *J. Phys. Chem.* **1993**, *97*, 3409−3417.
(38) Nose, S. *J. Chem. Phys.* **1989**, *81*, 511.
(39) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.