Calculation of the Binding Affinity of β-Secretase Inhibitors Using the Linear Interaction Energy Method

Brett A. Tounge and Charles H. Reynolds*

Johnson & Johnson Pharmaceutical Research and Development, L.L.C., P.O. Box 776, Welsh and McKean Roads, Spring House, Pennsylvania 19477-0776

Received November 15, 2002

It has been shown that the rate-limiting step in the production of β-amyloid peptide (Aβ) is the proteolytric cleavage of the membrane-bound β-amyloid precursor protein (APP) by β-secretase (BACE). Since the accumulation of Aβ has been implicated as one of the key events in the progression of Alzheimer's disease, BACE has become an important therapeutic target. Recently, two crystal structures of BACE cocrystallized with the inhibitors OM99-2 and OM00-3 were published by Tang and co-workers. In addition, the Ghosh group has published binding data on a series of inhibitors based on their initial lead, OM99-2. Using this set as a basis, we have developed a model for the binding affinity of these ligands to BACE using the linear interaction energy method. The best binding affinity model for the full set of ligands had a RMSD of 1.10 kcal/mol. The best model excluding the two charged ligands had a RMSD of 0.87 kcal/mol.

Introduction

The accumulation of β-amyloid peptide (Aβ) has been implicated as one of the key events in the progression of Alzheimer's disease.[1] This peptide is produced through the proteolytric cleavage of the membrane-bound β-amyloid precursor protein (APP) by β- and γ-secretases. Since β-secretase (β-site APP cleaving enzyme, BACE) has been shown to be the rate-limiting step in the production of Aβ in vivo,[2] it has become a target for drug discovery.[3−6] Tang and co-workers have determined the crystal structure of BACE complexed with two eight-residue peptides, OM99-2 (K_i = 1.6 nM) (Figure 1) and OM00-3 (K_i = 0.32 nM).[7,8] On the basis of these leads, the Ghosh group synthesized a series of related inhibitors and determined K_i values.[9]

We have used this series to investigate building a binding affinity model. The ability to predict ligand−protein binding affinities has obvious utility in drug discovery. As a result, several methods have been developed for addressing this difficult problem.[10] These techniques represent a tradeoff between computational time and accuracy. At one extreme are free-energy perturbation calculations .[11] These types of simulations are theoretically rigorous and as a result can produce quite accurate binding energy predictions. However, they entail extensive computer time and are not practical when even a moderate number of compounds need to be tested. On the other end of the spectrum are scoring functions. These functions are derived from empirical fits to an array of simplified energy terms that are meant to capture the key contributions to ligand−protein binding.[12−19] Because of these simplifications, the errors in the binding energies can be quite high, but these methods are extremely fast and thus well suited to high-volume screening of compound databases.

As a compromise between accuracy and computational speed, Åqvist and co-workers developed the linear interaction energy (LIE)[20] method, which has been successfully applied to many systems with errors on the order of 1 kcal/mol.[21−36] At the heart of this method is the assumption that the free energy of binding can be derived from considering only the two end points of the thermodynamic cycle of ligand binding. Simulations are carried out for the ligand free in solution and bound to a protein (still in the presence of solvent). In the original formulation of the method, molecular dynamics (MD) simulations were performed for each state to obtain the average (denoted by ⟨...⟩) intermolecular van der Waals (vdw) and electrostatic (elec) interactions (U). The binding free energy (ΔG_{bind}) was then derived using the following formula,

$$\Delta G_{bind} = \alpha\langle\Delta U_{vdw}\rangle + 0.5\langle\Delta U_{elec}\rangle \quad (1)$$

where the Δ term indicates the change in energy from the ligand free and bound states ($U_{bound} - U_{free}$). The van der Waals term was included to capture the nonelectrostatic contributions and was not meant to be a calculation of the actual energy contribution to ΔG. Instead, it is assumed that the nonpolar terms are correlated to the van der Waals interaction. As a result, the adjustable parameter, α, which is determined by fitting to experimental data, was included.

Since the publication of the original paper, Åqvist and others have explored several variants of the method. Much of this work has focused on investigating the transferability of the α and β parameters, as well as the need for adding a third parameter to the fit. For example, in later publications, Åqvist proposed a case-dependent β parameter (based on the charge and number of hydroxyl groups).[37] In addition, several other groups have made use of a third parameter in the fit, as first proposed by Jorgensen and co-workers,

$$\Delta G_{bind} = \alpha\langle\Delta U_{vdw}\rangle + \beta\langle\Delta U_{elec}\rangle + \gamma\Delta SASA \quad (2)$$

where ΔSASA is the change in the solvent accessible

* To whom correspondence should be addressed. Tel: (215) 628-5675. Fax: (215) 628−4985. E-mail: creynol1@prdus.jnj.com.

10.1021/jm020513b CCC: $25.00
Published on Web 04/26/2003

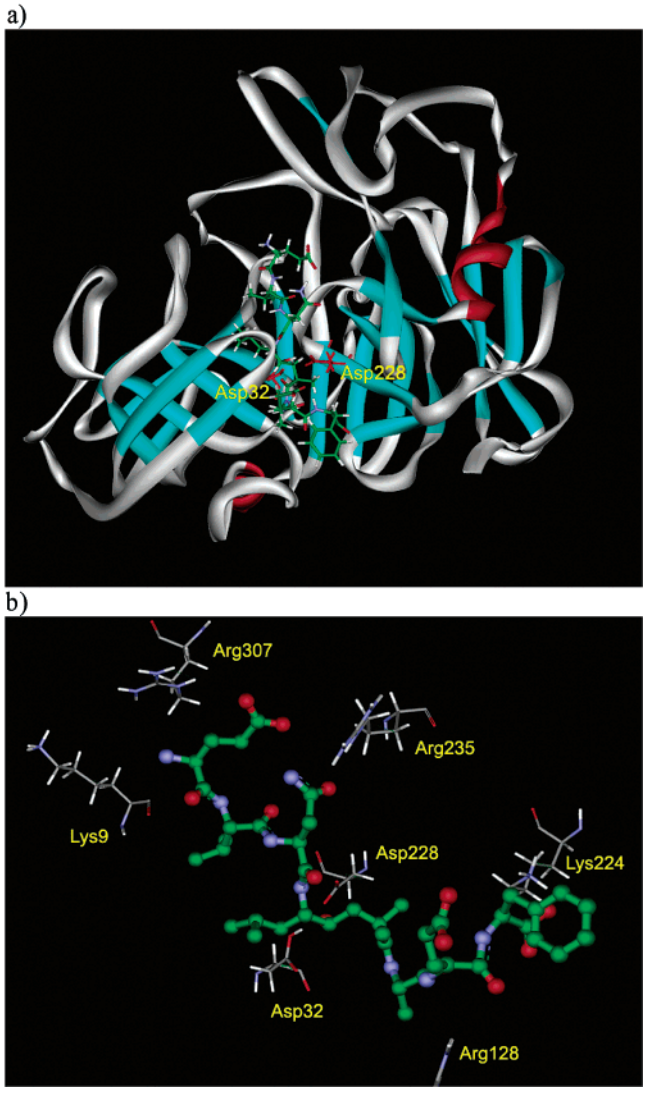


Figure 1. Crystal structure of OM99-2 bound to BACE (a). Figure b highlights some of the charged residues in the binding pocket.

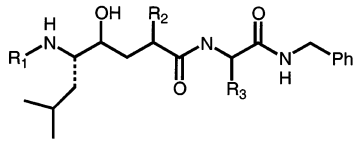
surface area.[38] A detailed statistical analysis of these fits, as well as other variants of the LIE method, has been reported by Wall et al.[39]

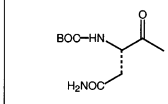
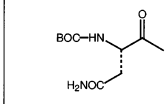
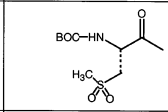
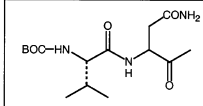
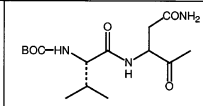
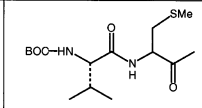
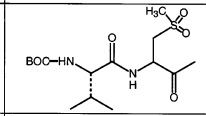
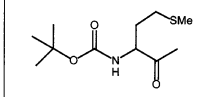
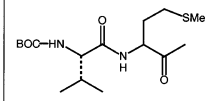
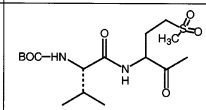
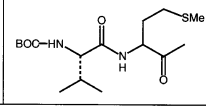
In this paper, we focus on the application of the LIE method to the set of BACE inhibitors outlined previously. Particular attention is paid to the effects of the protein charge state and residue-based electrostatic cutoff relative to the quality of the derived binding affinity model.

Computational Details

Simulations were performed on a set of 12 ligands that comprise a series of BACE inhibitors synthesized by Ghosh et al. (Table 1 and Figure 2).[9] These ligands were designed on the basis of the crystal structure of a potent BACE inhibitor, OM99-2 (eight-residue peptide, $K_i = 1.6$ nM). Since to a large extent they all preserve the backbone structure of the original peptide, the initial docked position of the ligands without X-ray structures was obtained by analogy to the OM99-2 bound conformation. Only two ligands in the series, OM00-3 and OM99-2, are charged (Figure 2). Experimental binding affinities in the form of K_i's have been published. This

Table 1. BACE Inhibitors Designed by Ghosh and Coworkers[9] [a]

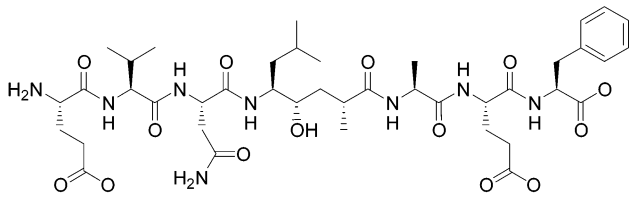


Comp.	R_1	R_2	R_3	K_i(nM)	ΔG (kcal/mol)
1		Me	Me	22423.0	-6.38
2		Me	$CHMe_2$	3134.0	-7.55
3		Me	$CHMe_2$	1129.0	-8.16
4		Me	Me	61.4	-9.90
5		Me	$CHMe_2$	5.9	-11.30
6		Me	$CHMe_2$	50.1	-10.02
7		Me	$CHMe_2$	9.4	-11.02
8		Me	$CHMe_2$	5808.0	-7.19
9		Me	$CHMe_2$	2.5	-11.81
10		Me	$CHMe_2$	8.0	-11.11
11		CH_2CHMe_2	$CHMe_2$	10491.0	45.11

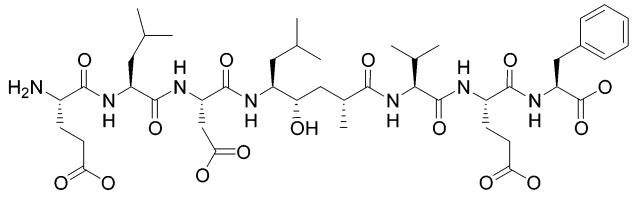
[a] This series was developed on the basis of the crystal structure of OM99-2 bound to β-secretase.

set covers a 6.7 kcal/mol range. One of the ligands from the series (compound **11**) was not used in building the model. However, its predicted binding affinity is discussed in the Results section.

The energy calculations were carried out using the Liaison package from Schrödinger Inc.[40] This implementation of LIE makes use of a surface-generalized Born (SGB) continuum solvation model,[41] a procedure that was first suggested by Zhou et al.[42] All charges are treated using the OPLS-AA force field.[43] In this implementation of the force field, no bond angle or stretch



OM99-2: K_i = 1.6 nM ΔG_{bind} = -12.06 kcal/mol



OM00-3: K_i = 0.32 nM ΔG_{bind} = -13.05 kcal/mol

Figure 2. Structures of peptide BACE inhibitors: OM99-2 and OM00-3.

parameters exist for the sulfone group. As a result, sulfonamide values were substituted for the missing sulfone parameters. The validity of these parameters has been confirmed by comparing bond distances and angles from B3LYP/6-31G*-optimized sulfone and sulfonamide groups (done by Schrödinger Inc.).[44] From these calculations, it was determined that no additional sulfone parameters were necessary.

Two methods were employed for obtaining the average LIE energies. In the first procedure, a conjugate gradient minimization was performed, followed by collection of the LIE energies. This procedure was carried out using a 10 and 15 Å residue-based cutoff as well as with no cutoff. In the second procedure, the conjugate gradient minimization was followed by a hybrid Monte Carlo (HMC) step that employed 20 ps of heating (before collecting the LIE energies), a sample target temperature of 300 K, and 30 ps of sampling for the LIE energies (total simulation time = 50 ps). For this set of HMC simulations, the time step was 0.002 ps, there were five MD steps per HMC cycle, and energies were sampled every 10 steps. Some representative plots of the average LIE energies versus time are shown in Figures 3 and 4 (the full set of plots is provided in the Supporting Information). For both the noncharged (Figure 3) and charged ligands (Figure 4), 50 ps is more than adequate to obtain converged energies. In most cases, converged energies are obtained with shorter HMC runs (20 ps total simulation time, 10 ps heating, and 10 ps of sampling). However, for consistency, 50 ps was used for each ligand. For all the simulations, protein residues beyond 10 Å from the binding pocket (as defined by ligand OM00-3) were frozen. Given the use of the continuum solvation model, the SASA term used in eq 2 is replaced by the cavity term of the continuum solvation model. The final equation used for fitting is

$$\Delta G_{bind} = \alpha\langle\Delta U_{vdw}\rangle + \beta\langle\Delta U_{elec}\rangle + \gamma\Delta cav \qquad (3)$$

Least-squares fits, based on singular value decomposition, to this equation were done using the Liaison package.

In this version of the Liaison package (FirstDiscovery v2.0), the SGB contribution to offsetting electrostatic interactions truncates at 12 Å. As a result, electrostatic interactions between atoms pairs that are greater than

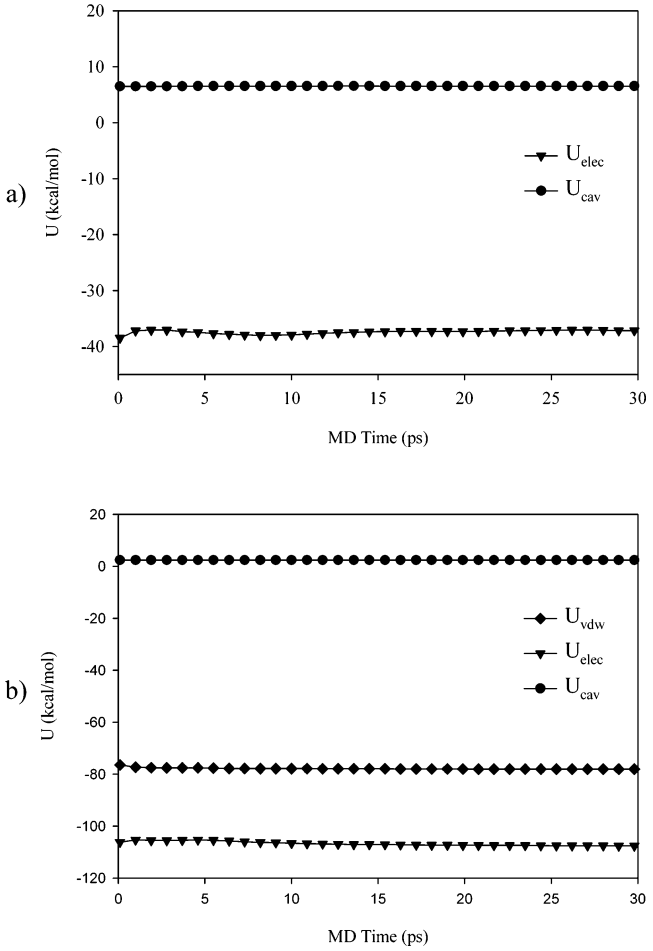


Figure 3. Average LIE energy terms for structure **6** versus MD time. These are the energies following a 20 ps MD equilibration time. Plot a is for the free state and plot b is for the bound state.

12 Å apart are not shielded by solvent. This fact, along with the emphasis placed by Åqvist on the importance of the charge state of the system, requires that careful attention be paid to which residues are charged.[45] For comparison, three different variants of charge neutralization were employed (using the protein preparation utility provided by Schrödinger Inc.). In method I, all residues within 9.6 Å of compound **1**, as well as residues that formed salt bridges (defined by a 3.5 Å cutoff), were left charged. Residues outside this range were then adjusted to make the overall charge of the protein neutral. This resulted in a +3 charge within the 9.6 Å shell around compound **1**. For method II, all residues outside of 12 Å from OM00-3 were neutralized and all residues within 6.5 Å of OM00-3 were charged. Residues between 6.5 and 12 Å of OM00-3 were then adjusted to give an overall neutral charge for the protein. For method III, only Asp228 and Arg235 were charged. In all cases, the protonation state of the catalytic aspartic acids (Asp228 and Asp32) was set so as to have a net charge of −1. This is consistent with studies of HIV-1 protease to which BACE has a similar pH rate profile.[46,47] The choice of which Asp to protonate was guided by the crystal structure of OM99-2. Given the position of the OM99-2 hydroxyl group, the most favorable hydrogen-bonding interactions are obtained by protonating Asp32. In this configuration, the OH of Asp32 can act as a hydrogen bond donor and the negatively

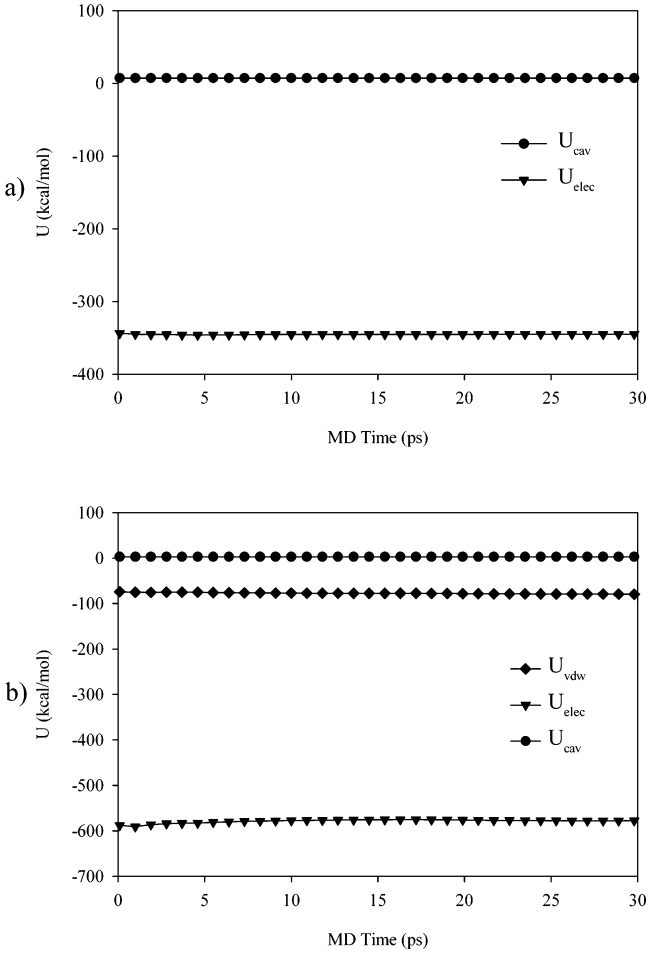


Figure 4. Average LIE energy terms for OM99-2 (charged ligand) versus MD time. These are the energies following a 20 ps MD equilibration time. Plot a is for the free state and plot b is for the bound state.

charged Asp228 can act as a hydrogen bond acceptor to the OM99-2 hydroxyl group. Of course, it is impossible to know the protonation state unambiguously. However, as long as a consistent protonation state is chosen, there should be little effect on the derived binding affinity model because all of the ligands used in this training set preserve the same binding motif at the catalytic site.

Results

As stated previously, two major factors that influence the energy terms in the LIE equation are the treatment of the protein charge state and the sampling method used to obtain the average energies. Much work has been published surrounding these two issues, so a careful investigation of the models as a function of these parameters was undertaken. Toward this goal, six LIE calculations were performed to isolate the different effects. A summary of the fits to eq 3 can be found in Table 2. Although the RMSD values for each model are comparable, some of the derived models have aspects that are not intuitive. In two out of the six models, the cavity term has a negative coefficient. Of the remaining four models, a significant deviation from the linear response approximation occurs. The β coefficient is much lower than the theoretical value of 0.5−0.33. The contribution of each of the simulation parameters to these fits is outlined below.

Table 2. Summary of the Fits to Eq 3 Using Different Protein Preparation Methods and Cutoffs for Nonbonded Interactions[a]

	method	α	β	γ	RMSD
A	protein prep. = method I sampling = minimization cutoff = 10 Å	0.041	0.033	1.372	1.492
B	protein prep. = method I sampling = minimization cutoff = 15 Å	0.229	0.060	−2.682	1.107
C	protein prep. = method I sampling = minimization cutoff = none	0.019	0.034	1.768	1.446
D	protein prep. = method I sampling = HMC cutoff = 15 Å	0.239	0.042	−2.422	1.101
E	protein prep. = method II sampling = minimization cutoff = 15 Å	0.055	0.015	1.240	1.511
F	protein prep. = method III sampling = minimization cutoff = 15 Å	0.052	0.015	1.287	1.474

[a] The best fit (based on RMSD) is obtained using protein preparation method I and a cutoff of 15 Å.

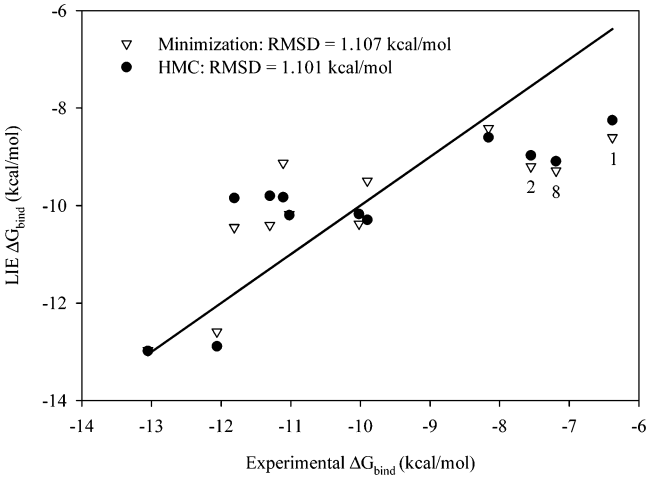


Figure 5. Comparison of the experimental versus calculated ΔG_{bind} for two different sampling methods: minimization and HMC. For both procedures, protein preparation method I and a 15 Å cutoff were used. The compound numbers for ligands with high errors are noted on the plot.

Sampling Method. Using protein preparation method I, we used two different sampling procedures in determining the LIE average energies (Table 3). For this set of ligands it was found that the conformation space search provided by the HMC method did not produce a significantly better fit to the experimental binding energies than simple minimization (1.101 kcal/mol for HMC versus 1.107 kcal/mol for minimization; see Figure 5). This is due in large part to the level of accuracy of the starting bound conformations. As was noted in the Computational Details, all the ligands to some extent preserve the backbone structure of OM99-2. Therefore, we were able to prepare starting conformations that had a high probability of being near the global minimum structures for the bound ligands.

Electrostatic Term. The difficulty in treating the electrostatic energy term has been well noted in the literature and is not particular to the LIE method. The variability of this term as a function of the protein preparation method is illustrated in Table 4 for the BACE inhibitors. The calculation of this term can be broken down into two factors: the protein charge

Table 3. LIE Terms for the Two Different Sampling Methods[a]

ligand	Δvdw^{MIN}	Δvdw^{HMC}	$\Delta elec^{MIN}$	$\Delta elec^{HMC}$	Δcav^{MIN}	Δcav^{HMC}
1	−71.848	−69.016	−30.149	−26.006	−3.819	−3.891
2	−74.805	−71.983	−36.498	−33.934	−3.974	−3.997
3	−75.069	−69.321	−20.485	−30.240	−3.950	−4.160
4	−81.666	−81.394	−34.956	−43.140	−4.042	−4.156
5	−85.08	−80.797	−48.162	−34.310	−4.160	−4.319
6	−85.665	−78.067	−47.266	−45.393	−4.155	−4.194
7	−85.835	−80.091	−38.372	−41.868	−4.363	−4.386
8	−77.379	−69.956	−34.013	−38.174	−4.064	−4.121
9	−85.669	−80.743	−44.712	−34.565	−4.336	−4.431
10	−87.732	−80.758	−14.346	−34.192	−4.384	−4.446
OM99-2	−78.661	−79.414	−90.73	−89.740	−4.549	−4.437
OM00-3	−72.859	−73.425	−167.330	−198.350	−5.14	−5.296

[a] Both simulations were done using protein preparation method I and a 15 Å cutoff. The terms from the hybrid Monte Carlo and minimization runs are indicated with HMC and MIN, respectively.

Table 4. Electrostatic Term for Each of the LIE Calculations[a]

	ΔU_{elec} (kcal/mol) for each method					
ligand	A	B	C	D	E	F
1	−28.665	−30.149	−30.281	−26.006	−33.518	−32.882
2	−29.762	−36.498	−24.185	−33.934	−25.503	−31.877
3	−15.263	−20.485	−20.450	−30.240	−22.478	−12.667
4	−42.372	−34.956	−35.707	−43.140	−34.767	−35.832
5	−33.104	−48.162	−33.714	−34.310	−34.779	−37.626
6	−32.835	−47.266	−35.538	−45.393	−37.505	−37.044
7	−24.968	−38.372	−14.888	−41.868	−26.594	−2.057
8	−23.554	−34.013	−23.567	−38.174	−27.010	−23.674
9	−30.196	−44.712	−36.524	−34.565	−38.457	−38.171
10	−24.524	−14.346	−20.332	−34.192	−10.835	−15.550
OM99-2	−13.490	−90.730	−25.220	−89.740	2.640	−12.430
OM00-3	−69.460	−167.330	−78.54	−198.35	−92.05	−90.82

[a] The details for each method are listed in Table 2.

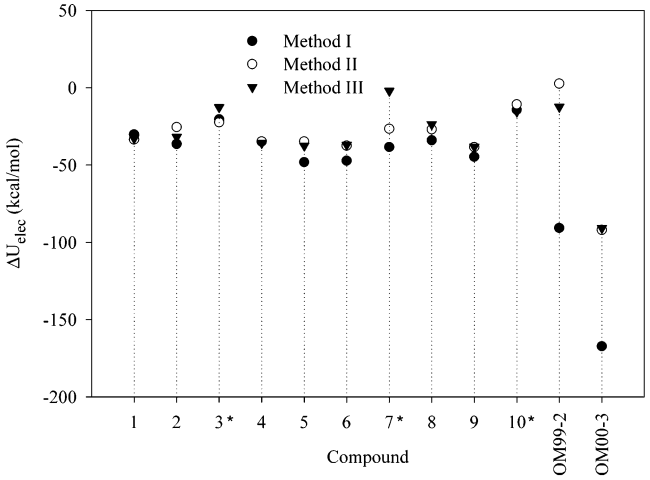


Figure 6. Protein preparation method I was used to perform three different LIE calculations where the nonbonded cutoff was 10 Å, 15 Å, and no cutoff. The resulting ΔU_{elec} is plotted for each ligand. Compounds containing a sulfone group are marked with an asterisk.

preparation method and the cutoff used for the nonbonded interactions.

The effect of the nonbonded cutoff is illustrated in Figure 6. For these studies, method I was used for the protein preparation. Not surprisingly, the most significant changes in the electrostatic terms occur for the charged ligands. As a result of the 12 Å solvent cutoff from the continuum solvation model as implemented in the Schrödinger package, the charged groups of OM99-2 and OM00-3 are seeing unshielded (by solvent) charges when the nonbonded cutoff is increased. The next most significant change occurs for one of the sulfone-containing ligands (ligand **7**). The sulfur atom of this group is assigned a partial charge of 1.374, and the oxygens are assigned a partial charge of −0.687 (OPLS-AA force field; see Computational Details). Because of the lack of shielding, these atoms are seeing large changes as the nonbonded cutoff is increased.

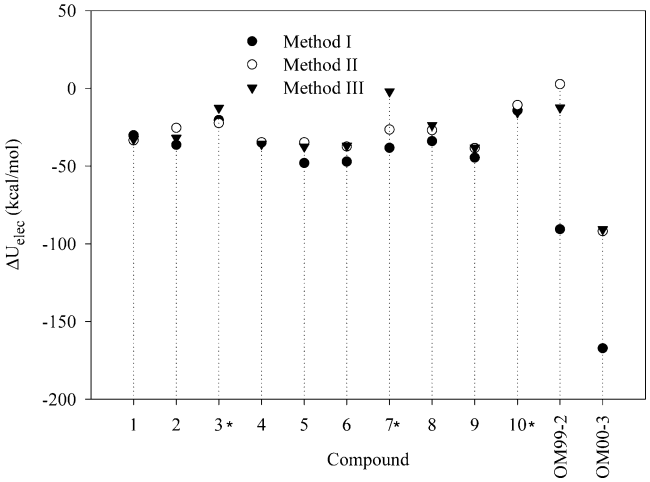


Figure 7. For this set of simulations, the cutoff for the nonbonded interactions was held at 15 Å. The protein preparation method was then varied. Compounds containing a sulfone group are marked with an asterisk.

The effect of the charged state of the protein is illustrated in Figure 7. In this case, a nonbonded cutoff of 15 Å was used for each of the simulations. Here again we see that the ligands most affected by the protein preparation are charged or have a group with a large point charge. This example illustrates that overall charge neutrality is not the only issue to consider when preparing the protein. The location of the groups that remain charged also has a significant effect. In practice, it can be difficult a priori to determine which of the charged residues are the most significant in terms of their contribution. For the simulations run here, method I allows for the most positively charged groups to stabilize the −3 charge of OM00-3.

As was noted in the Computational Details section, one of the ligands from Ghosh's work was left out of the models discussed above. According to the experimental results, the substitution of a methyl group with CH_2-$CHMe_2$ at the R_2 position results in a change in K_i from 2.5 to 10491 nM (compounds **9** and **11** respectively). However, no energy penalty is found in the LIE calculations (Table 5). As a result, compound **11** is predicted to have almost the same activity (predicted ΔG_{bind} =

Table 5. Comparison of LIE Terms for Compounds **9** and **11**[a]

LIE term (kcal/mol)	**9**	**11**
ΔU_{vdw}	−80.743	−84.738
ΔU_{elec}	−34.565	−37.782
ΔU_{cav}	−4.431	−4.686

[a] These structures differ only in the R_2 position. Compound **9** contains a methyl group, and compound **11** has a CH_2CHMe_2 group. Although their experimental K_i's are drastically different, there is very little change in the LIE terms. Both simulations were carried out using HMC sampling, a 15 Å cutoff, and protein preparation method I.

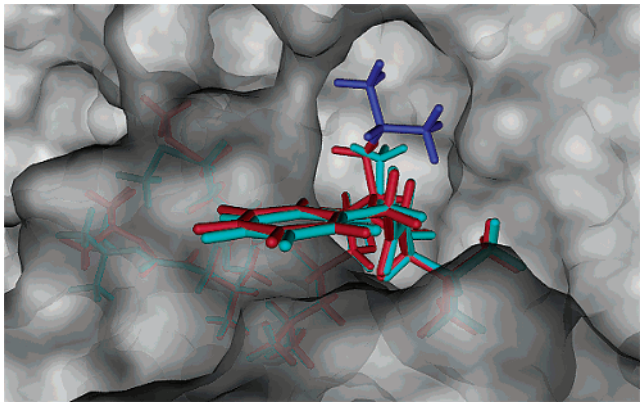


Figure 8. Binding pocket of BACE with compound **9** (cyan) and compound **11** (red). According to the experimental data, there is a significant loss in activity when the methyl group in compound **9** is substituted with a CH_2CHMe_2 group (blue) in compound **11** at the R_2 position. However, from the LIE calculations there is no penalty for the additional bulk on compound **11**. A visual inspection reveals no steric hindrance for this group.

−10.481 kcal/mol) as compound **9** (fitted ΔG_{bind} = −9.987 kcal/mol). It is not clear whether this is a failure of the LIE method or if there is a problem with the experimental K_i. As can be seen in Figure 8 (binding pocket with both compounds **9** and **11**), there is significant overlap in the constant part of each ligand and no obvious steric hindrance for the additional bulk of the $CHMe_2$ group. In fact, the computed van der Waals energy is slightly more favorable for compound **11** because of this additional group. In this case, more sampling would not help because we would never capture a higher-energy structure.

Cavity Term. Another issue that has been debated in the literature is the need for a third parameter in the LIE equation. Two issues must be considered when adding this third parameter. First, given that many of the systems studied to date using the LIE method have had a small training set, the possibility of over-fitting the data exists with the addition of a third parameter. Second, one must be cautious when applying a linear regression method that the correlation among terms is kept to a minimum. For the set of Ghosh ligands, the correlation coefficient between the cavity and electrostatic terms is 0.89. As a result, addition of the third parameter does not add significant information to the model. The fit using just the van der Waals and electrostatic term yields:

$$\Delta G \text{ (kcal/mol)} = 0.114\langle\Delta U_{vdw}\rangle + 0.0246\langle\Delta U_{elec}\rangle \quad (4)$$

with a RMSD of 1.202 kcal/mol (Figure 9). This is essentially equivalent to the RMSD of 1.101 kcal/mol

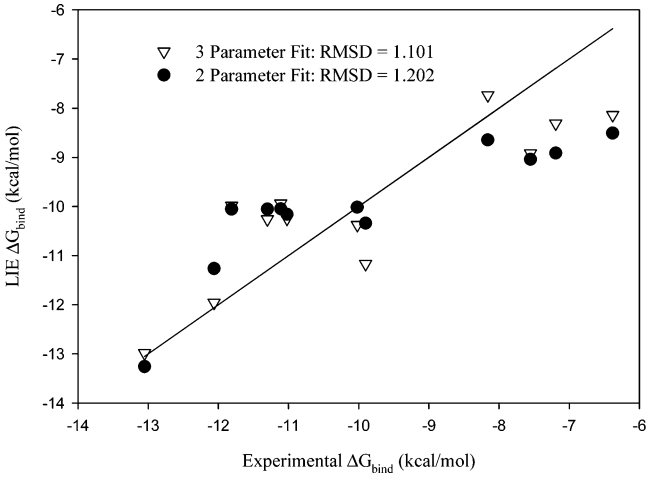


Figure 9. Comparison of the experimental versus calculated ΔG_{bind} for two variants of the LIE equation. Fit 1 uses the three-parameter fit of eq 3, whereas Fit 2 is a fit using only the van der Waals and electrostatic terms.

Table 6. Predicted Free Energies of Binding for Simulation Method D Using Both a Two- and Three-Parameter Fit

	ΔG_{bind} (kcal/mol)		
ligand	2-parameter fit	3-parameter fit	experimental
1	−8.50815	−8.139	−6.38
2	−9.04178	−8.920	−7.55
3	−8.64728	−7.736	−8.16
4	−10.3415	−11.170	−9.90
5	−10.0557	−10.261	−11.30
6	−10.0178	−10.377	−10.02
7	−10.1616	−10.246	−11.02
8	−8.91523	−8.313	−7.19
9	−10.0559	−9.987	−11.81
10	−10.0484	−9.939	−11.11
OM99-2	−11.2644	−11.960	−12.06
OM00-3	−13.259	−12.983	−13.05

found for the fit to all three terms (See Table 6 for predicted values). This issue is discussed in more detail by Wall.[39]

Charged Ligands. The literature for LIE would suggest caution when using a training set of ligands that differ in their overall charged state.[20] For comparison, a LIE model was also derived using only the noncharged ligands (Figure 10). Using this more homogeneous set, a slightly worse fit was obtained for the HMC run (RMSD = 1.150 kcal/mol). However the fit for the minimization only run was improved, with the RMSD dropping from 1.107 to 0.872 kcal/mol. It is encouraging to note that the model containing the two charged ligands is comparable to the model containing only neutral ligands. Since it is possible to construct a good model with the charged ligands present, most of the analysis in this study was carried out for the full set.

Discussion

We have been able to successfully model a set of BACE ligands that differ significantly in terms of their overall charge state. Using this mixed set, a RMSD of 1.101 kcal/mol was obtained. The robustness of this model is further illustrated by the RMSD of 1.355 kcal/mol determined by jackknife cross-validation (Table 7). Using this method we find that no one compound is having a disproportionate contribution to the overall binding affinity model. However, a closer investigation

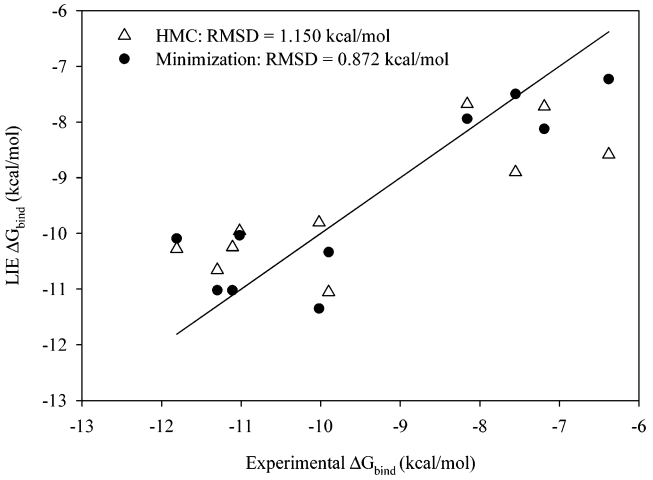


Figure 10. Experimental versus calculated ΔG_{bind}, excluding the charged ligands (OM99-2 and OM00-3). These simulations use protein preparation method I and a 15 Å cutoff.

Table 7. LIE and Cross Validation (Jackknife)-Predicted Free Energies of Binding for Simulation Method D[a]

	ΔG_{bind} (kcal/mol)		
ligand	experimental	LIE	jackknife
1	−6.38	−8.25	−8.416
2	−7.55	−8.973	−9.057
3	−8.16	−8.605	−7.471
4	−9.90	−10.297	−11.850
5	−11.30	−9.801	−10.110
6	−10.02	−10.174	−10.433
7	−11.02	−10.197	−10.158
8	−7.19	−9.094	−8.601
9	−11.81	−9.848	−9.753
10	−11.11	−9.833	−9.784
OM99-2	−12.06	−12.888	−11.919
OM00-3	−13.05	−12.170	−12.050
	RMSD	1.101	1.355

[a] The jackknife value for each ligand is obtained by predicting its binding affinity from the parameters obtained by a fit to the other ligands in the set.

of the trends of the individual terms of the LIE model versus the binding energy does illustrate some areas where the model merits further scrutiny.

Plots of the correlations between the experimental ΔG_{bind} and each term in the LIE equation are shown in Figure 11 for the best fit (based on RMSD) model that includes the charged ligands. One of the features highlighted by these plots is that the two charged ligands, OM99-2 and OM00-3, exhibit large deviations from the linear trend as defined by the rest of the (neutral) ligands in the set. This is particularly true for the electrostatic term. The charged ligands both have small van der Waals energies relative to the other ligands, given their binding affinities. However, these are offset by favorable electrostatics. As a result, there are large and compensating deviations from the linear relationship between the ΔG_{bind} and the van der Waals and electrostatic terms. This raises concerns about the validity of assuming a constant scaling factor (β) for the electrostatic term when these highly charged ligands are included. Many factors might be responsible for this observation, including a breakdown in the linear response assumption for such large changes in the charged state of the ligands, limitations in the underlying methods for computing the electrostatic or nonbonded energy terms, or all of these factors. In any event, it is

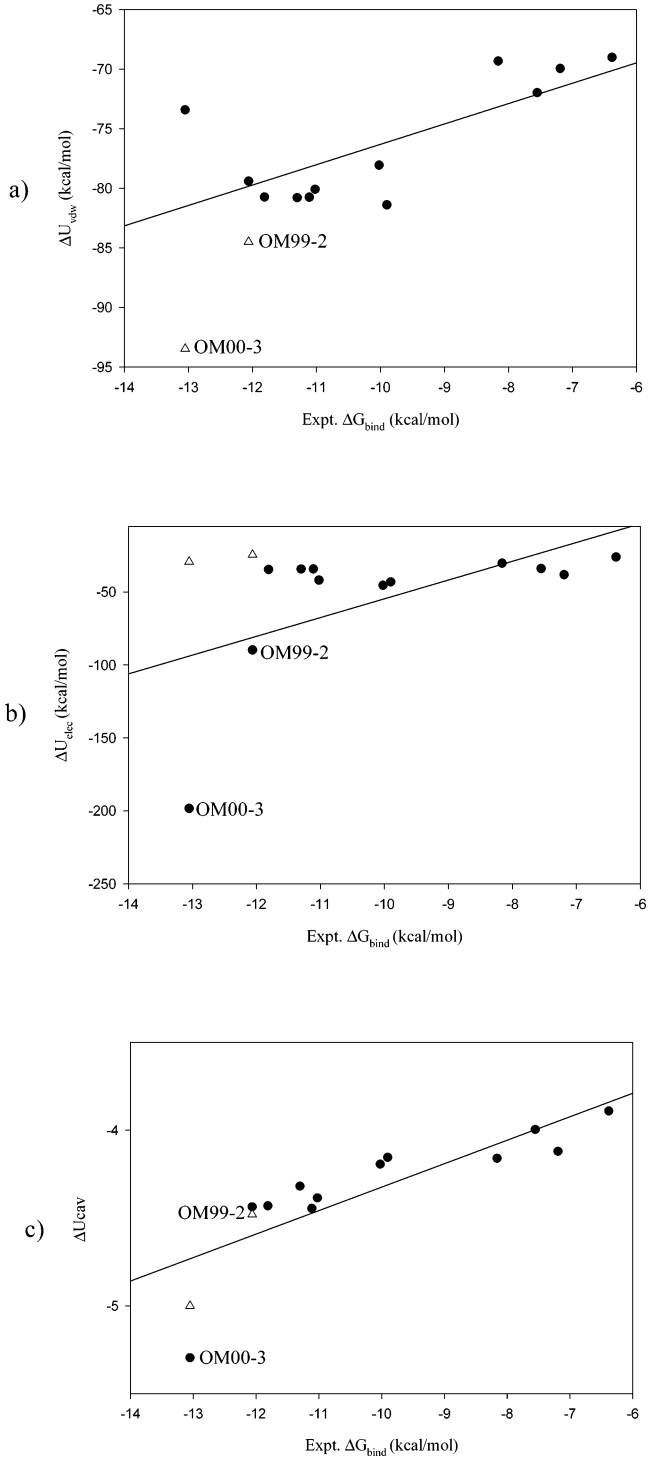


Figure 11. Individual trends for the LIE energies (van der Waals (a), electrostatic (b), cavity (c)) versus the experimental ΔG_{bind}. These energies are for simulation method D. The triangles represent the values for OM00-3 and OM99-2 when run as neutral ligands.

clear that balancing the subtle interactions between electrostatic and van der Waals terms poses a serious challenge. For example, if OM00-3 is treated as a neutral ligand, the van der Waals energy falls near the line in Figure 11a for the other ligands. When treated as the charged ion, the van der Waals energy is exceptionally low (Table 8). This means that the sampled configuration is forced into a geometrically poor fit by the more favorable electrostatics. These results certainly highlight the complex balance of forces (i.e., nonbonded

Table 8. Change in the van der Waals and Electrostatic Energies for OM00-3 and OM99-2 as a Function of the Ligand Charge State (Simulation Method D)

	OM00-3		OM99-2	
	ΔU_{vdw}[a]	ΔU_{elec}[a]	ΔU_{vdw}[a]	ΔU_{elec}[a]
neutral	−93.482	−29.238	−84.492	−34.404
charged	−73.425	−198.350	−79.414	−89.740

[a] In kcal/mol.

and electrostatic) that must be modeled to predict the total ΔG_{bind}. They also imply that the electrostatic energy plays a dominant role in this interplay. It is possible, but far from certain, that a more theoretically rigorous treatment of the electrostatics would improve this situation.

Conclusions

A predictive model for the binding affinities of the full series β-secretase inhibitors was obtained using the LIE method (RMSD = 1.101 kcal/mol). By comparison, the best model obtained where the two charged ligands were omitted is only somewhat better with a RMSD of 0.872 kcal/mol. The model for the full set represents the best RMSD obtained for a series of fits using different sampling techniques and protein treatments. For this set of inhibitors it was found that there was little improvement in the model when HMC sampling was used as opposed to just minimizing the docked ligands and then collecting the LIE energy terms. This is probably a result of the knowledge of the initial docked position from the two crystal structures. Using this sampling technique, we also investigated different combinations of protein charged states and electrostatic cutoffs. The best model was obtained using a charged residue shell of 9.6 Å around the ligand in combination with a 15 Å residue-based electrostatic cutoff. The robustness of the model is well-illustrated by the 1.355 kcal/mol RMSD obtained using the leave-one-out cross validation technique. This indicates that the fit is not biased by any particular ligand. Given the robustness of the model, it should be valuable in assessing new BACE inhibitors.

Acknowledgment. Brett Tounge thanks the Johnson & Johnson Corporate Office of Science and Technology for support of his postdoctoral fellowship. In addition, Brett Tounge and Charles Reynolds thank Allen Reitz and Christopher Creighton of the Spring House CNS team and Hege Beard from Schrödinger for their input.

Supporting Information Available: Plots of the average energy versus HMC step for each of the LIE terms for all the ligands. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Selkoe, D. J. Translating Cell Biology into Therapeutic Advances in Alzheimer's Disease. *Nature* **1999**, *399A*, A23-A31.
(2) Sinha, S.; Lieberburg, I. Cellular Mechanisms of β-Amyloid Production and Secretion. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11049−11053.
(3) Dingwall, C. Spotlight on BACE: The Secretases as Targets for Treatment in Alzheimer's Disease. *J. Clin. Invest.* **2001**, *108*, 1243−1246.
(4) Fang, L. Y.; John, V. PCT Int. Appl.; (Elan Pharmaceuticals, Inc.: USA). WO, 0202506, 2002; p 434.
(5) Olson, R. E.; Thompson, L. A. Secretase Inhibitors as Therapeutics for Alzheimer's Disease. *Annu. Rep. Med. Chem.* **2000**, *35*, 31−40.
(6) Schenk, D.; Games, D.; Seubert, P. Potential Treatment Opportunities for Alzheimer's Disease through Inhibition of Secretases and A.β. Immunization. *J. Mol. Spectrosc.* **2001**, *17*, 259−267.
(7) Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A. K.; Zhang, X. C.; Tang, J. Structure of the Protease Domain of Memapsin 2 (β-Secretase) Complexed with Inhibitor. *Science (Washington, D. C.)* **2000**, *290*, 150−153.
(8) Hong, L.; Turner, R. T., III; Koelsch, G.; Shin, D.; Ghosh, A. K.; Tang, J. Crystal Structure of Memapsin 2 (β-Secretase) in Complex with an Inhibitor OM00-3. *Biochemistry* **2002**, *41*, 10963−10967.
(9) Ghosh, A. K.; Bilcer, G.; Harwood, C.; Kawahama, R.; Shin, D.; Hussain, K. A.; Hong, L.; Loy, J. A.; Nguyen, C.; Koelsch, G.; Ermolieff, J.; Tang, J. Structure-Based Design: Potent Inhibitors of Human Brain Memapsin 2 (β-secretase). *J. Med. Chem.* **2001**, *44*, 2865−2868.
(10) Lazaridis, T. Binding Affinity and Specificity from Computational Studies. *Curr. Org. Chem.* **2002**, *6*, 1319−1332.
(11) Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(12) Holloway, M. K. A Priori Prediction of Ligand Affinity by Energy Minimization. *Perspect. Drug Discov.* **1998**, *9/10/11*, 63−84.
(13) Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.
(14) Bohm, H.-J.; Stahl, M. Rapid Empirical Scoring Functions in Virtual Screening Applications. *Med. Chem. Res.* **1999**, *9*, 445−462.
(15) Muegge, I. A Knowledge-Based Scoring Function for Protein−Ligand Interactions: Probing the Reference State. *Perspect. Drug Discovery Des.* **2000**, *20*, 99−114.
(16) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical Scoring Functions. II. The Testing of an Empirical Scoring Function for the Prediction of Ligand−Receptor Binding Affinities and the Use of Bayesian Regression To Improve the Quality of the Model. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 503−519.
(17) Pearlman, D. A.; Charifson, P. S. Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(18) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(19) Wang, R.; Wang, S. How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422−1426.
(20) Aqvist, J.; Marelius, J. The Linear Interaction Energy Method for Predicting Ligand Binding Free Energies. *Comb. Chem. High Throughput Screening* **2001**, *4*, 613−626.
(21) Brandsdal, B. O.; Aqvist, J.; Smalas, A. O. Computational Analysis of Binding of P1 Variants to Trypsin. *Protein Sci.* **2001**, *10*, 1584−1595.
(22) Chen, J.; Wang, R.; Taussig, M.; Houk, K. N. Quantitative Calculations of Antibody−Antigen Binding: Steroid-DB3 Binding Energies by the Linear Interaction Energy Method. *J. Org. Chem.* **2001**, *66*, 3021−3026.
(23) Dey, I. Exploring the Interaction of Some *N*-Benzyloxycarbonyl-L-phenyl alanyl-L-alanine Ketones and Bovine Spleen Cathepsin B by Molecular Modeling and Binding Free Energy Calculation. *J. Biomol. Struct. Dyn.* **1999**, *16*, 891−900.
(24) Graffner-Nordberg, M.; Kolmodin, K.; Qvist, J.; Queener, S. F.; Hallberg, A. Design, Synthesis, Computational Prediction, and Biological Evaluation of Ester Soft Drugs as Inhibitors of Dihydrofolate Reductase from *Pneumocystis carinii*. *J. Med. Chem.* **2001**, *44*, 2391−2402.
(25) Hou, T. J.; Zhang, W.; Xu, X. J. Binding Affinities for a Series of Selective Inhibitors of Gelatinase-A Using Molecular Dynamics with a Linear Interaction Energy Approach. *J. Phys. Chem. B* **2001**, *105*, 5304−5315.
(26) Hou, T.-J.; Zhang, W.; Xu, X.-J. Binding Free Energy Calculations for MMP2-Hydroxamate Complexes. *Huaxue Xuebao* **2002**, *60*, 221−227.
(27) Jones-Hertzog, D. K.; Jorgensen, W. L. Binding Affinities for Sulfonamide Inhibitors with Human Thrombin Using Monte Carlo Simulations with a Linear Response Method. *J. Med. Chem.* **1997**, *40*, 1539−1549.
(28) Lamb, M. L.; Tirado-Rives, J.; Jorgensen, W. L. Estimation of the Binding Affinities of FKBP12 Inhibitors Using a Linear Response Method. *Bioorg. Med. Chem.* **1999**, *7*, 851−860.
(29) Ljungberg, K. B.; Marelius, J.; Musil, D.; Svensson, P.; Norden, B.; Aqvist, J. Computational Modelling of Inhibitor Binding to Human Thrombin. *Eur. J. Pharm. Sci.* **2001**, *12*, 441−446.

(30) Luzhkov, V. B.; Aqvist, J. Mechanisms of Tetraethylammonium Ion Block in the KcsA Potassium Channel. *FEBS Lett.* **2001**, *495*, 191−196.
(31) Marelius, J.; Graffner-Nordberg, M.; Hansson, T.; Hallberg, A.; Aqvist, J. Computation of Affinity and Selectivity: Binding of 2,4-Diaminopteridine and 2,4-Diaminoquinazoline Inhibitors to Dihydrofolate Reductases. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 119−131.
(32) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 393−407.
(33) Smith, R. H., Jr.; Jorgensen, W. L.; Tirado-Rives, J.; Lamb, M. L.; Janssen, P. A. J.; Michejda, C. J.; Smith, M. B. K. Prediction of Binding Affinities for TIBO Inhibitors of HIV-1 Reverse Transcriptase Using Monte Carlo Simulations in a Linear Response Method. *J. Med. Chem.* **1998**, *41*, 5272−5286.
(34) Wang, J.; Dixon, R.; Kollman, P. A. Ranking Ligand Binding Affinities with Avidin: A Molecular Dynamics-Based Interaction Energy Study. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 69−81.
(35) Wang, W.; Wang, J.; Kollman, P. A. What Determines the van der Waals Coefficient β in the LIE (Linear Interaction Energy) Method To Estimate Binding Free Energies Using Molecular Dynamics Simulations? *Proteins: Struct., Funct., Genet.* **1999**, *34*, 395−402.
(36) Wesolowski, S. S.; Jorgensen, W. L. Estimation of Binding Affinities for Celecoxib Analogues with COX-2 via Monte Carlo-Extended Linear Response. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 267−270.
(37) Hansson, T.; Marelius, J.; Aqvist, J. Ligand Binding Affinity Prediction by Linear Interaction Energy Methods. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27−35.
(38) Carlson, H. A.; Jorgensen, W. L. An Extended Linear Response Method for Determining Free Energies of Hydration. *J. Phys. Chem.* **1995**, *99*, 10667−10673.
(39) Wall, I. D.; Leach, A. R.; Salt, D. W.; Ford, M. G.; Essex, J. W. Binding Constants of Neuraminidase Inhibitors: An Investigation of the Linear Interaction Energy Method. *J. Med. Chem.* **1999**, *42*, 5142−5152.
(40) *First Discovery*, 2.0 ed.; Schrödinger, Inc.: Portland, OR, 2001.
(41) Ghosh, A.; Rapp, C. S.; Friesner, R. A. Generalized Born Model Based on a Surface Integral Formulation. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.
(42) Zhou, R.; Friesner, R. A.; Ghosh, A.; Rizzo, R. C.; Jorgensen, W. L.; Levy, R. M. New Linear Interaction Method for Binding Affinity Calculations Using a Continuum Solvent Model. *J. Phys. Chem. B* **2001**, *105*, 10388−10397.
(43) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(44) Beard, H. Schrödinger, Inc., Portland, OR. Personal communication 2002.
(45) Aqvist, J. Calculation of Absolute Binding Free Energies for Charged Ligands and Effects of Long-Range Electrostatic Interactions. *J. Comput. Chem.* **1996**, *17*, 1587−1597.
(46) Gruninger-Leitch, F.; Schlatter, D.; Kung, E.; Nelbock, P.; Dobeli, H. Substrate and Inhibitor Profile of BACE β-Secretase) and Comparison with other Mammalian Aspartic Proteases. *J. Biol. Chem.* **2002**, *277*, 4687−4693.
(47) Hyland, L. J.; Tomaszek, T. A. J.; Meek, T. D. Human Immunodeficiency Virus-1 Protease. 2. Use of pH Rate Studies and Solvent Kinetic Isotope Effects To Elucidate Details of Chemical Mechanism. *Biochemistry* **1991**, *30*, 8454−8463.

JM020513B