JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

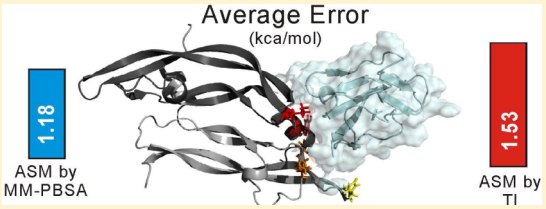
Computational Alanine Scanning Mutagenesis: MM-PBSA vs TI

Sílvia A. Martins,† Marta A. S. Perez,† Irina S. Moreira, Sérgio F. Sousa, M. J. Ramos, and P. A. Fernandes*

REQUIMTE/Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

Supporting Information

ABSTRACT: Understanding protein−protein association and being able to determine the crucial residues responsible for their association (hot-spots) is a key issue with huge practical applications such as rational drug design and protein engineering. A variety of computational methods exist to detect hot-spots residues, but the development of a fast and accurate quantitative alanine scanning mutagenesis (ASM) continues to be crucial. Using four protein−protein complexes, we have compared a variation of the standard computational ASM protocol developed at our group, based on the Molecular Mechanics/Poisson−Boltzmann Surface Area (MM-PBSA) approach, against Thermodynamic Integration (TI), a well-known and accurate but computationally expensive method. To compare the efficiency and the accuracy of the two methods, we have calculated the protein−protein binding free energy differences upon alanine mutation of interfacial residues ($\Delta\Delta G_{bind}$). In relation to the experimental $\Delta\Delta G_{bind}$ values, the average error obtained with TI was 1.53 kcal/mol, while the ASM protocol resulted in an average error of 1.18 kcal/mol. The results demonstrate that the much faster ASM protocol gives results at the same level of accuracy as the TI method but at a fraction of the computational time required to run TI. This ASM protocol is therefore a strong and efficient alternative to the systematic evaluation of protein−protein interfaces, involving hundreds of amino acid residues in search of hot-spots.



I. INTRODUCTION

Proteins participate in almost every level of cell function. However, often they do not accomplish their function on their own, they need to associate with other molecules, namely other proteins, in order to fulfill their biological functions.[1] Understanding protein−protein association and being able to determine the crucial residues responsible for their association has been a subject of intense research in the last decades. Bogan and Thorn demonstrated that only a few residues in a protein−protein interface are responsible for the binding: the hot-spots.[2] These are defined as residues that upon alanine mutation generate a binding free energy difference ($\Delta\Delta G_{bind}$) higher than 2.0 kcal/mol.[2] The correct detection of these residues is a key issue with huge practical application such as rational drug design and protein engineering.[3] Alanine scanning mutagenesis (ASM) has been widely applied to the characterization of these interfaces. However, experimental ASM is a costly and time-consuming task, which urged the need for fast and accurate theoretical methods. A huge amount of algorithms of increasing complexity have been employed to address the binding energy between biological molecules. They can be divided essentially into three types: empirical functions or simple physical methods that use knowledge-based simplified models to evaluate complex association; fully atomistic methods that estimate the binding free energy as a result of mutating the residues of the interacting molecules; or, more recently, feature-based approaches.[4] The feature-based approaches tend to be more qualitative than quantitative.[5−12] Therefore, an atomistic and accurate quantitative ASM method is still crucial to detect hot-spots.

Free energy is probably the most important quantity in thermodynamics and one of the central topics in biophysics. Nevertheless, for many relevant systems with local minimum energy configurations separated by energy barriers, efficient and accurate calculation of this property is still a big challenge in computational chemistry. Thermodynamic integration (TI) is the key choice to perform accurate calculations of the binding strength of protein complexes. This rigorous method yields accurate free energy differences relying on equilibrium sampling of an entire transformation path, from an initial to a final state. It is implemented numerically and utilizes a thermodynamic cycle and the fact that the free energy is a state function. However, as sufficient statistical sampling must be carried out, the use of TI turns out to be computationally very intensive and it is therefore limited in the screening of a large number of structural perturbations. Another methodological approach, which has become more attractive in the past few years for estimating binding free energies of protein−protein complexes, is the MM-PBSA (Molecular Mechanics/Poisson−Boltzmann Surface Area) method.[13−15] This method is a fully atomistic approach that combines molecular mechanics and continuum solvent. A few years ago, we developed a simple computational protocol that relies on the MM-PBSA approach but combines the use of different dielectric constants when different residues are mutated to alanine. The conformational sampling, the relaxation, and reorganization due to the mutation for an alanine are not explicitly included in the MM-PBSA formalism.

Received: January 15, 2013
Published: February 12, 2013

dx.doi.org/10.1021/ct4000372 | J. Chem. Theory Comput. 2013, 9, 1311−1319

Therefore, the scaling of the macroscopic parameter (internal dielectric constant) to larger values when larger reorganizations are expected mimics these effects. Using a set of three internal dielectric constants exclusively characteristic of the mutated amino acid (2 for the nonpolar amino acids, 3 for the polar residues, and 4 for the charged amino acids plus histidine), it was possible to increase the agreement with the experimental results for the $\Delta\Delta G_{bind}$ values. To test our ASM protocol against such an accurate method as TI, we chose to apply them both to four distinct proteic complexes: (i) Vascular Endothelial Growth Factor and FLT-1 Receptor (PDBID:1FLT);[16] (ii) Barnase and Barnstar (PBDID:1BRS);[17] (iii) Bacterial cell division ZipA and Ftsz (PDBID:1F47[2]);[18] and (iv) Igg1 Kappa D1.3 FV and Hen Egg white lysozyme (PDBID:1VFB[4]).[19] These systems were selected based on the existence of experimental binding free energy ($\Delta\Delta G_{bind}$) values for the interfacial residues upon alanine mutations and their dissimilar properties in terms of size, chemical and physical character. Their biological importance and typical interfaces makes them a perfect data set.

II. METHODOLOGY

A. System Setup. In our four reference systems, protonation states of the different residues were determined using the PDB2PQR server at http://kryptonite.nbcr.net/pdb2pqr/[20] with the PROPKA methodology.[21−23]

B. Alanine Scanning Mutagenesis. *a. Molecular Dynamics Simulations.* The MD simulations were performed using the AMBER9 package[24] with the Duan et al. force field.[25] Two different simulations were performed: one in an implicit solvent using the Generalized Born (GB) solvent[26] and another using TIP3P explicit water molecules. Each complex was solvated by explicit waters that extended 10 Å from any edge of the box to the protein atoms. Counter ions were added to the boxes to neutralize the system. In each of the simulations, the system was initially minimized to remove bad contacts using the steepest descent algorithm followed by conjugated gradient. The systems were then subjected to 2 ns of heating (in NVT ensemble) in which the temperature was gradually raised to 300 K, followed by 6 ns runs in the NPT ensemble. The Langevin[27,28] thermostat was used, and the electrostatics interactions were calculated using the particle mesh Ewald (PME) method.[29] Both lengths involving hydrogen atoms were constrained using the SHAKE algorithm.[30] The equations of motion were integrated with a 2 fs time-step and the nonbonded interactions were truncated with a 16 Å and a 10 Å cutoff, in the GB and in explicit solvent simulations, respectively. The final structures resulting not only from both minimizations (in explicit and implicit solvent) but also from the MD simulations trajectories were subsequently subjected to alanine mutation.

b. Alanine Scanning Mutagenesis Protocol. The MM-PBSA (Molecular Mechanics Poisson−Boltzmann Surface Area) script[15] integrated into the AMBER9 package[24] was used to calculate the binding free energy difference ($\Delta\Delta G_{bind}$) upon alanine mutation. It combines a continuum approach to model solvent interactions with an MM-based approach to atomistically model protein−protein interactions. The protein structures used to calculate the binding free energy may come either from a MD simulation or just from energy minimization of the X-ray structures. This provides speed and accuracy and has been quite used in the last years.[4,13,15,31−39] The MM-PBSA approach first developed by Massova et al.[15] was adapted by Moreira et al.[4] to implement an accurate ASM protocol. In the case of geometry-optimized structures, the mutant complexes are generated by a single truncation of the mutated side chain, replacing Cα with a hydrogen atom and setting the Cα-H direction to that of the former Cα-Cβ. In the case of the structures generated by MD simulations, a total of 320 snapshots of the complexes were extracted in the last 1 ns of the run. The $\Delta\Delta G_{bind}$ is written as the difference between the mutant and wild type complexes defined as

$$\Delta\Delta G_{bind} = \Delta G_{bind}^{mut} - \Delta G_{bind}^{wt} \quad (1)$$

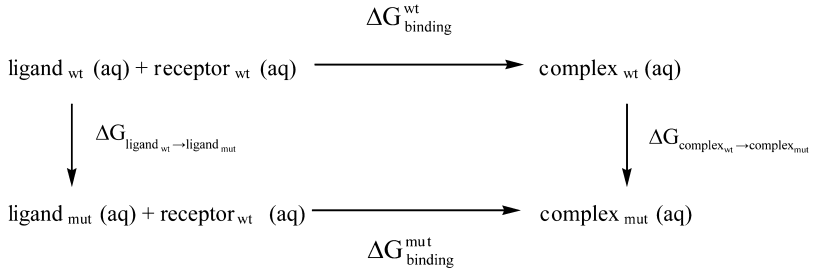
Typical contributions to the free energy include the internal energy (bond, dihedral, and angle), the electrostatic and the van der Waals interactions, the free energy of polar solvation, the free energy of nonpolar solvation, and the entropic contribution:

$$G_{molecule} = E_{int\ ernal} + E_{electrostatic} + E_{vdW} + G_{polar\ solvation} + G_{nonpolar\ solvation} - TS \quad (2)$$

For the calculations of relative free energies between closely related complexes, it is assumed that the entropic contributions are negligible as it essentially cancel each other in eq 2.[13] The first three terms of eq 2 were calculated with no cutoff. The $G_{polar\ solvation}$ term was calculated by solving the Poisson−Boltzmann equation with the software DELPHI.[40,41] In this continuum method, the protein is modeled as a dielectric continuum of low polarizability embedded in a dielectric medium of high polarizability. We have used a set of values for the DELPHI parameters that in a previous study have constituted a good compromise between accuracy and computational speed.[42] Therefore, we used a value of 2.5 grids/Å for scale (the reciprocal of the grid spacing); a value of 0.001 kT/c for the convergence criterion; a 90% for the fill of the grid box; and the Coulombic method to set the potentials at the boundaries of the finite-difference grid. The dielectric boundary was taken as the molecular surface defined by a 1.4 Å probe sphere and by spheres centered on each atom with radii taken from the Parse[43] vdW radii parameter set. The key aspect of our ASM protocol is the use of three dielectric constants (with the value 2 for nonpolar residues, 3 for polar residues, and 4 for charged residues plus histidine) to mimic the expected rearrangement upon alanine mutation. It is important to highlight that we have used only one trajectory for the computational energy analysis, as it has been shown to give the best results.[4] Side-chain reorientation was implicitly included in the formalism by raising the internal dielectric constant. The nonpolar contribution to the solvation free energy due to van der Waals interactions between the solute and the solvent was modeled as a term dependent on the solvent accessible surface area (SASA) of the molecule. It was estimated to be 0.00542 X SASA+0.92 using the molsurf program developed by Mike Connolly.[44] As a systematic mutation of residues on protein−protein interfaces (PPI) is a fastidious and time-consuming methodological approach, we have recently developed a VMD[45] plugin (http://compbiochem.org/Software/compasm/Home.html).[46] This plugin has a friendly graphical interface and was used in this work.

C. Thermodynamic Integration. The thermodynamic integration (TI) method allows for the calculation of the difference in free energy between two given states. Equation 3 can be derived directly from the configuration integral. From

Scheme 1. Thermodynamic Cycle for Calculating the Binding Free Energy Difference between the Wild Type Residues and the Mutant Residues in the Four Complexes Considered[a]



[a] $\Delta G^{wt}_{binding}$ and $\Delta G^{mut}_{binding}$ are binding free energies for the wild type and the mutant respectively.

this equation the free energy between two states of a given system can be obtained using the coupling parameter (λ) approach. The coupling parameter varies from 0 to 1 corresponding respectively to the initial A and final B states. Basically, the free energy difference, $\Delta G_{A\to B}$, is the integral from 0 to 1 of the expectation value of $\partial V(\lambda)/\partial\lambda$, where V is the potential energy. The integral in eq 3 may be evaluated numerically using a number of discrete λ points.

$$\Delta G_{A\to B} = \int_0^1 \left(\frac{\partial V(\lambda)}{\partial \lambda}\right)_\lambda d\lambda \tag{3}$$

The TI method was used to calculate the difference in the free energy of binding, upon mutation of an interface residue to an alanine ($\Delta\Delta G_{bind}$) in order to test the effectiveness of the ASM protocol. The $\Delta\Delta G_{bind}$ was evaluated using the thermodynamic cycle shown in Scheme 1.

From the thermodynamic cycle, we get eqs 5 and 6:

$$\Delta\Delta G_{bind} = \Delta G^{mut}_{bind} - \Delta G^{wt}_{bind} \tag{5}$$

$$\Delta\Delta G_{bind} = \Delta G_{complex_{wt}\to complex_{mut}} - \Delta G_{ligand_{wt}\to ligand_{mut}} \tag{6}$$

Two different transformations need to be simulated, wild type to mutant in the ligand, $\Delta G_{ligand_{wt}\to ligand_{mut}}$ and wild type to mutant in the complex $\Delta G_{complex_{wt}\to complex_{mut}}$. For reasons of simulation stability, these two transformations have been divided into three substeps each: first, the atomic partial charges on the side chain atoms were removed (ΔG^1); second, the van der Waals (vdW) potentials and radii were transformed from the wt values into the alanine residues (ΔG^2); and finally, the side chain had its atomic partial charges switched on to their alanine values (ΔG^3). This was done because having a nonzero charge on an atom while the vdW interactions with its surroundings are getting weaker can lead to well-known simulation instabilities. Also, for reasons of simulation stability, we have used softcore potentials in substep 2, which are modified Lennard-Jones potentials that prevent simulation instabilities due to the truncation of the potential to small energy values for small or zero radius. We note that it is impossible to directly assign a Coulombic or vdW partitioning to the total free energy as these are both path dependent.

We can now express $\Delta G_{ligand_{wt}\to ligand_{mut}}$, $\Delta G_{complex_{wt}\to complex_{mut}}$, and $\Delta\Delta G_{bind}$ as a function of ΔG^1, ΔG^2, and ΔG^3, and we get eqs 7, 8, and 9:

$$\Delta G_{ligand_{wt}\to ligand_{mut}} = \Delta G^1_{ligand} + \Delta G^2_{ligand} + \Delta G^3_{ligand} \tag{7}$$

$$\Delta G_{complex_{wt}\to complex_{mut}} = \Delta G^1_{complex} + \Delta G^2_{complex} + \Delta G^3_{complex} \tag{8}$$

$$\Delta\Delta G_{bind} = (\Delta G^1_{complex} + \Delta G^2_{complex} + \Delta G^3_{complex}) - (\Delta G^1_{ligand} + \Delta G^2_{ligand} + \Delta G^3_{ligand}) \tag{9}$$

We have computed the free energy of each substep of each transformation with the AMBER10 software.[47] Each substep was performed in explicit solvent and under periodic boundary conditions with nine λ values (0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90). For each λ value running in each substep, we have carried out 500 steps of steepest descent minimization, a 50 ps density equilibration run, and a 200 ps NPT production run. The total simulation time to mutate just one wild type residue into an alanine was 13 500 ps. Free energy derivatives ($\partial V/\partial\lambda$) were collected independently for each λ from the production run. A time step of 1 fs is used together with the SHAKE algorithm. Ewald sums with a 9 Å cutoff in the real part, isotropic pressure scaling, and a Langevin type thermostat to maintain the temperature at 300 K were also used. Each system was centered in a cubic box of water with a minimum distance of 12 Å between any protein atom and the box side. The standard amino acid residues were accounted for by the use of the Duan et al. AMBER force field.[25] The TIP3P water model was used.

It is important to notice here that the conditions used in the ASM and TI studies differ in some particular aspects (e.g., the time step used and the size of the water buffer). For a perfect comparison, it would be, in principle, preferable to keep the same conditions in both studies. However, our objectives here were more ambitious. We wanted to show that our computational ASM protocol could compete in terms of accuracy with Thermodynamic Integration even when a more rigorous TI protocol was employed. For this reason, we have used a smaller, more rigorous time step in TI together with a larger water buffer in TI.

III. RESULTS AND DISCUSSION

Experimental ASM is long, laborious, and costly. An important advantage of computer simulations over experiments is not only to provide faster estimates of the binding free energy difference but also to enhance our understanding of the nature of complex formation in terms of the biophysical features of the process, because they add molecular insight into the macroscopic properties measured therein. Computational ASM is a tool that, if well used, can assist experimental ASM by making it more capable and more profitable. A reliable computational ASM protocol would allow minimizing the number of experimental

assays carried out, because it would identify the residues that are most probably hot-spots and the residues that will almost surely be null-spots. Therefore, it is important to find an atomistic and accurate quantitative ASM protocol capable of reproducing the experimental mutagenesis values.

To test our already established ASM protocol against such an accurate method as TI, we have calculated the $\Delta\Delta G_{bind}$ with both methodological alternatives for the four systems. Experimental $\Delta\Delta G_{bind}$ values were used as reference values.[18,48−52] We have also analyzed advantages and disadvantages for both methods as well as their range of applicability and limitations, their expected performance, and their precision and accuracy.

A. Data Set. Our data set comprises four complexes with various chemical and physical characteristics. The complex VEGF:FLT-1 (Figure 1) has a high biological importance and a relatively small interfacial area. To decrease the computational time involved in the calculations with both methods, we have used the VEGF dimer and only one FLT-1 monomer. The complex Barnase−Barnstar is a well know complex with a very charged interface; the complex between the bacterial cell division ZipA and Fts has a small, hydrophobic interface, and the complex between the Igg1 Kappa D1.3 FV and the Hen Egg white lysozyme has the largest of the interfaces under study. The data set of 22 residues has 45% of hot-spots and 55% of null-spots, and within these groups, the ΔΔGbind has a large range: 2.29−6.00 and 0.00−1.80 kcal/mol, respectively. To better understand the chemical composition of the data set, we can divided it in three groups: charged (Asp and Glu, His, Lys and Arg), polar (Ser, Thr, Asn, Gln, Tyr), and nonpolar (Val, Ile, Leu, Met, Phe, Trp). The percentages within our data set are 36%, 32%, and 32%, respectively. So, it represents a perfect data set test for the comparison and the development of computational methods.

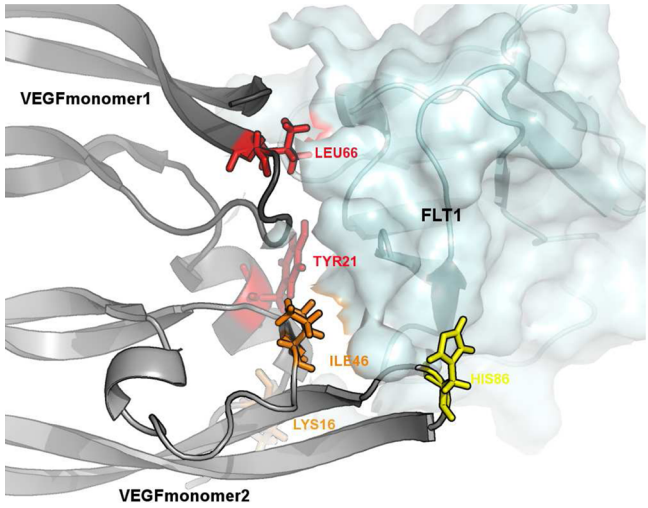


Figure 1. Representation of the interface between the VEGF dimer and the FLT-1 monomer highlighting Tyr21 and Leu66 in red (hot-spots), Lys16 and Ile46 in orange, and His186 in yellow (PDBID:1FLT). VEGF monomer1 is represented in black tube, VEGF monomer2 in gray tube, the FLT-1 receptor in surface.

B. ASM Protocol. The final structures from both the minimizations (in explicit and implicit solvent) as well as the trajectories from the MD simulations (in explicit and implicit solvent) were subjected to mutation and subsequent calculation of $\Delta\Delta G_{bind}$. In Supporting Information, it can be seen that in both cases (implicit and explicit) the 320 snapshots were extracted from the last nanosecond of the fully equilibrated part of the MD simulations.

The $\Delta\Delta G_{bind}$ between the wild-type residues and the alanine mutant variants obtained with the ASM protocol are presented in Table 1. The uncertainties of the calculated values (standard deviations) and the comparisons with the experimental values are also presented in Table 1.

Different variations of the ASM protocol are shown in Table 1: (1) implicit solvent molecular dynamics (320 structures); (2) implicit solvent with molecular mechanics minimization only (1 structure); (3) explicit solvent with molecular dynamics (320 structures); (4) explicit solvent with molecular mechanics minimization only (1 structure).

The results from the protocols involving only molecular mechanics (MM) show the worst agreement with the experimental values, particularly those from the MM in implicit solvent, with an average deviation from the experimental value of 2.51 kcal/mol and a maximum difference of 7.25 kcal/mol. The standard deviation of the mean, defined as $\sigma/\sqrt{n}$, where n is the number of snapshots and σ is the standard deviation between snapshots, is not presented for the MM results because such results are based on one single structure, the MM optimized structure. As expected, an extremely fast minimization with a single structure subjected to the ASM protocol is not sufficient to obtain accurate results.

From both MD, implicit and explicit solvent, it is worth noting that there are differences in the precision (related with standard deviation, the ability of the measurement to be consistently reproduced), accuracy (how close a result comes to the experimental value, average error), and reliability (related with maximum error, the ability to generate the same result). There are significant differences in the individual results, with the standard deviation ranging from 0.21 to 1.20 kcal/mol in implicit solvent and from 0.2 to 1.47 kcal/mol in explicit solvent. The results from the implicit solvent are, hence, more precise. The largest differences from the experimental values are obtained in explicit solvent, with a mean error of 1.77 kcal/mol and a maximum error of 5.89 kcal/mol while the mean and maximum errors from the MD in implicit solvent are significantly lower, 1.18 and 4.69 kcal/mol, respectively. The results from implicit solvent are therefore also in this case more accurate.

The use of the implicit solvent to calculate $\Delta\Delta G_{bind}$ leads to a better agreement with the experimental data. The preference for the implicit solvent over the explicit solvent can be justified by several reasons, namely the smaller simulation time necessary compared to that of the explicit solvent method, the more complete exploration of the conformational space due to the lack of the viscous damping forces of the water, the reduced lengthy equilibration of water compared to that of the explicit water simulation, and the easier interpretation of the results since the water degrees of freedom are absent. Additionally, the ASM protocol was optimized for the use of an MD trajectory of the wild-type system in implicit solvent. Moreover, the MD in implicit solvent for the complex VEGF:FLT-1 used as reference is 4.2 times faster than the MD in explicit solvent in our cluster. The ASM protocol in implicit solvent has been used with success in the study of several biological systems, including the IgG1 streptococcal protein G (C2 fragment) complex,[32] the FTase complex,[53] and the antibody HyHEL-10[34] with the antibody FVD1.3,[36] and the

Table 1. Differences in Binding Free Energies between the Wild-Type Residues and the Alanine Mutant Variants Obtained with ASM Protocol[a]

			MM-PBSA									
			Implicit Solvent					Explicit Solvent				
			MD			MM		MD			MM	
	mutation	$\Delta\Delta G_{exp}$ [b]	$\Delta\Delta G_{bind}$	SD	$\lvert\Delta\Delta G_{MM\text{-}PBSA} - \Delta\Delta G_{exp}\rvert$	$\Delta\Delta G_{bind}$	$\lvert\Delta\Delta G_{MM\text{-}PBSA} - \Delta\Delta G_{exp}\rvert$	$\Delta\Delta G$	SD	$\lvert\Delta\Delta G_{MM\text{-}PBSA} - \Delta\Delta G_{exp}\rvert$	$\Delta\Delta G_{bind}$	$\lvert\Delta\Delta G_{MM\text{-}PBSA} - \Delta\Delta G_{exp}\rvert$
1FLT	Lys16Ala	0.35	0.03	0.74	0.32	−1.27	1.62	0.57	0.99	0.22	0.75	0.40
	Tyr21Ala	2.85	2.14	1.00	0.71	−1.47	4.32	3.92	1.30	1.07	1.09	1.76
	Leu66Ala	2.28	1.37	1.20	0.91	−1.18	3.46	1.98	1.47	0.30	0.62	1.66
	Ile46Ala	0.82	0.02	1.20	0.80	−1.54	2.36	0.01	0.64	0.81	−0.17	0.99
	His86Ala	0.00	0.06	0.80	0.06	−0.75	0.75	−2.01	1.16	2.01	0.70	0.70
1BRS	Arg59Ala	5.2	6.68	0.57	1.48	12.45	7.25	11.09	0.36	5.89	10.53	5.33
	Arg87Ala	5.5	10.19	0.89	4.69	4.19	1.31	10.47	0.97	4.97	8.53	3.03
	His102Ala	6	5.71	0.58	0.29	6.68	0.68	6.83	0.36	0.83	6.68	0.68
	Tyr29Ala	3.4	5.68	0.60	2.28	8.21	4.81	4.80	1.07	1.40	8.21	4.81
	Thr42Ala	1.8	1.86	0.36	0.06	0.38	1.42	1.37	0.38	0.43	0.38	1.42
1F47	Phe9Ala	2.44	3.42	0.48	0.98	3.70	1.26	4.23	0.29	1.79	4.14	1.70
	Leu10Ala	2.29	2.95	0.47	0.66	1.14	1.15	2.71	1.09	0.42	2.74	0.45
	Asp2Ala	0.69	−1.55	0.51	2.24	0.08	0.61	−1.67	0.29	2.36	−2.28	2.97
	Tyr3Ala	0.86	3.20	0.42	2.34	4.74	3.88	5.87	1.28	5.01	3.02	2.16
	Leu4Ala	0.92	2.18	0.48	1.26	−0.29	1.21	2.18	0.29	1.26	−0.11	1.03
	Asp5Ala	1.73	−0.61	0.51	2.34	0.38	1.35	0.54	0.30	1.19	−0.42	2.15
1VFB	Trp92Ala	1.71	2.17	0.31	0.46	7.22	5.51	2.17	0.20	0.46	6.68	4.97
	Tyr101Ala	>4.0	3.80	0.87	>0.2	6.40	>2.40	5.05	1.01	>1.05	6.40	>2.4
	Val120Ala	0.9	1.33	0.34	0.43	2.62	1.72	1.12	0.54	0.22	0.55	0.35
	Gln121Ala	2.9	3.85	0.21	0.95	9.50	6.60	7.34	0.78	4.44	10.56	7.66
	Ser93Ala	0.11	0.22	0.42	0.11	−0.15	0.26	−0.21	0.22	0.32	−0.46	0.57
	Arg125Ala	1.8	3.11	0.75	1.31	2.89	1.09	3.67	0.50	1.87	4.53	2.73
	mean			⟨0.62⟩	⟨1.18⟩		⟨2.51⟩		⟨0.70⟩	⟨1.77⟩		⟨2.26⟩
	max.			1.20	4.69		7.25		1.47	5.89		7.66

[a]The uncertainties of the free energy differences and comparisons with the experimental values are also included. All values are in kcal/mol. [b]$\Delta\Delta G_{bind}$ is the difference in the free energy of binding; $\Delta\Delta G_{exp}$ is the experimental $\Delta\Delta G_{bind}$; $\Delta\Delta G_{MM\text{-}PBSA}$ is the $\Delta\Delta G_{bind}$ obtained with the ASM protocol; SD is the standard deviation; $\lvert\Delta\Delta G_{MM\text{-}PBSA}\text{-}\Delta\Delta G_{exp}\rvert$ is the absolute difference between the theoretical and experimental values; MM is molecular mechanics and MD is molecular dynamics.

dx.doi.org/10.1021/ct400037z | J. Chem. Theory Comput. 2013, 9, 1311−1319

Table 2. Differences in Binding Free Energies between the Wild-Type Residues and the Alanine Mutant Variants Obtained with Thermodynamic Integration[a]

	mutation	$\Delta\Delta G_{exp}$ [b] kcal/mol	$\Delta\Delta G_{TI}$ kcal/mol	SD	$\lvert\Delta\Delta G_{TI} - \Delta\Delta G_{exp}\rvert$
1FLT	Lys16Ala	0.35	1.28	0.74	0.93
	Tyr21Ala	2.85	2.55	0.48	0.30
	Leu66Ala	2.28	2.45	0.45	0.17
	Ile46Ala	0.82	−0.23	0.45	1.05
	His86Ala	0.00	0.09	0.51	0.09
1BRS	Arg59Ala	5.2	4.87	0.71	0.33
	Arg87Ala	5.5	8.23	0.76	2.70
	His102Ala	6.0	3.40	0.48	2.60
	Tyr29Ala	3.4	5.13	0.52	1.73
	Thr42Ala	1.8	2.96	0.41	1.16
1F47	Phe9Ala	2.44	5.62	0.44	3.18
	Leu10Ala	2.29	5.18	0.43	2.89
	Asp2Ala	0.69	0.67	0.64	0.02
	Tyr3Ala	0.86	5.95	0.50	5.09
	Leu4Ala	0.92	−1.82	0.46	2.74
	Asp5Ala	1.73	2.15	0.64	0.42
1VFB	Trp92Ala	1.71	0.93	0.55	0.78
	Tyr101Ala	>4.0	−1.01	0.56	>1.01
	Val120Ala	0.9	−0.95	0.46	1.85
	Gln121Ala	2.9	4.90	0.50	2.00
	Ser93Ala	0.11	−0.11	0.38	0.22
	Arg125Ala	1.8	3.66	0.73	1.86
	mean			⟨0.54⟩	1.53
	max			⟨0.76⟩	5.09

[a]Also shown are the uncertainties of the free energy differences and the comparisons with the experimental values. All values are given in kcal/mol. [b]$\Delta\Delta G_{exp}$ is the $\Delta\Delta G_{bind}$ experimental; $\Delta\Delta G_{TI}$ is the $\Delta\Delta G_{bind}$ obtained with TI; SD is the standard deviation error; $\lvert\Delta\Delta G_{TI} - \Delta\Delta G_{exp}\rvert$ is the absolute difference between the theoretical and experimental values.

MDM2-P53 complex.[37] Furthermore, in previous benchmarking studies against experimental data, it has been shown to have an overall success of 82% in identifying hot spots and to yield a mean unsigned error of around 0.8 kcal/mol.[4,33] The following section analyses the accuracy of the more computationally demanding thermodynamic integration method.

C. Thermodynamic Integration Method. In Table 2, we present the binding free energy differences as calculated by TI and the respective RMS, the correspondent uncertainty, and the comparison of the $\Delta\Delta G_{bind}$ with the experimental values.

To improve the results, more simulations could be added at different λ points (this is indeed one of the strong points of TI, you can add as many additional data points as you want to refine your result without having to redo the initial calculations), more production time could be used (for better convergence and more complete conformational sampling), or even more sophisticated numerical integration schemes could be used (we have used the trapezoidal rule to integrate numerically). The used protocol takes more than 15 days in an eight processor machine of our computer cluster for just one mutation, and the mentioned improvements would lead to a further increase of the simulation time. Relative to the experimental value, the largest differences obtained are for Tyr3 and Phe9, both for 1F47 with deviations of 5.09 and 3.18 kcal/mol, respectively. The $\Delta\Delta G_{bind}$ of the other residues tested are even closer to the experimental value with deviations of 0.02 kcal/mol, 0.09 and 0.17 kcal/mol, for Asp2 (1F47), His86, and Leu66 (1FLT), respectively.

D. MM-PBSA vs TI. TI is one of the most accurate methods to compute free energies. (Free Energy Perturbation is equally efficient and the difference mainly pertains to the formula used for evaluating the free energy). In the past decade, the computational ASM method has been shown to yield particularly accurate, precise, and reliable results. The main question that we try to answer in this study is "How competitive, in terms of accuracy and computational time, is the computational ASM protocol, in relation to TI, in calculating the change in the free energy of binding, upon mutation of an interfacial residue to an alanine?"

For that purpose, we compared the results of both methods in Table 3.

From Table 3, we can conclude that both methods are capable of predicting the experimental mutagenesis results. As far as the differences between calculated and experimental values are concerned, $\lvert\Delta\Delta G_{calc} - \Delta\Delta G_{exp}\rvert$ ranges from 0.06 to 4.69 with the ASM protocol and from 0.02 to 5.09 kcal/mol with TI. The average of $\lvert\Delta\Delta G_{calc} - \Delta\Delta G_{exp}\rvert$ for the 22 residues tested is 1.18 kcal/mol with the ASM protocol and a little higher (1.53 kcal/mol) with TI. ASM method is an atomistic quantitative computational method, capable of reproducing the experimental mutagenesis values.

There are several points from the methodological point of view that differ between both methods as far as alanine mutagenesis is concerned. TI is a computationally demanding methodology that produces reliable, although huge amount of data to be analyzed. On the other side, the ASM protocol produces much less data, using MM-PBSA a much faster methodology and a considerably easier technique. These limitations are particularly relevant when the interfaces are large, because the computational time with TI grows linearly with the number of mutations. For ASM, the most computationally demanding part of the calculation is the initial

Table 3. TI vs ASM in the Study of VEGF: FLT-1 Interface Residues[a]

			TI			MM-PBSA implicit solvent MD			
	mutation	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{bind}$	SD	$\lvert\Delta\Delta G_{calc} - \Delta\Delta G_{exp}\rvert$	$\Delta\Delta G_{bind}$	SD	$\lvert\Delta\Delta G_{calc} - \Delta\Delta G_{exp}\rvert$	$\lvert\Delta\Delta G_{ASM} - \Delta\Delta G_{TI}\rvert$
1FLT-1	Lys16Ala	0.35	1.28	0.74	0.93	0.03	0.74	0.32	1.25
	Tyr21Ala	2.85	2.55	0.48	0.30	2.14	1.02	0.71	0.41
	Leu66Ala	2.28	2.45	0.45	0.17	1.37	1.15	0.91	1.08
	Ile46Ala	0.82	−0.23	0.45	1.05	0.02	1.20	0.80	0.25
	His86Ala	0.00	0.09	0.51	0.09	0.06	0.82	0.06	0.03
1BRS	Arg59Ala	5.2	4.87	0.71	0.33	6.68	0.57	1.48	1.81
	Arg87Ala	5.5	8.23	0.76	2.70	10.19	0.89	4.69	1.96
	His102Ala	6.0	3.40	0.48	2.60	5.71	0.58	0.29	3.40
	Tyr29Ala	3.4	5.13	0.52	1.73	5.68	0.60	2.28	0.55
	Thr42Ala	1.8	2.96	0.41	1.16	1.86	0.36	0.06	1.10
1F47	Phe9Ala	2.44	5.62	0.44	3.18	3.42	0.48	0.98	2.20
	Leu10Ala	2.29	5.18	0.43	2.89	2.95	0.47	0.66	2.23
	Asp2Ala	0.69	0.67	0.64	0.02	−1.55	0.51	2.24	2.22
	Tyr3Ala	0.86	5.95	0.50	5.09	3.20	0.42	2.34	2.74
	Leu4Ala	0.92	−1.82	0.46	2.74	2.18	0.48	1.26	4.00
	Asp5Ala	1.73	2.15	0.64	0.42	−0.61	0.51	2.34	2.76
1VFB	Trp92Ala	1.71	0.93	0.55	0.78	2.17	0.31	0.46	1.24
	Tyr101Ala	>4.0	2.99	0.56	>1.01	3.80	0.87	>0.2	0.81
	Val120Ala	0.9	−0.95	0.46	1.85	1.33	0.34	0.43	2.28
	Gln121Ala	2.9	4.90	0.50	2.00	3.85	0.21	0.95	1.05
	Ser93Ala	0.11	−0.11	0.38	0.22	0.22	0.42	0.11	0.33
	Arg125Ala	1.8	3.66	0.73	1.86	3.11	0.75	1.31	0.55
	mean			0.54	1.53		0.62	1.18	1.56
	max.			0.76	5.09		1.20	4.69	4.00

[a] $\Delta\Delta G_{bind}$ is the difference in the free energy of binding; $\Delta\Delta G_{exp}$ is the $\Delta\Delta G_{bind}$ experimental; $\Delta\Delta G_{calc}$ is the $\Delta\Delta G_{bind}^{wt\rightarrow mut}$ obtained with ASM protocol; SD is the standard deviation error; $\lvert\Delta\Delta G_{calc} - \Delta\Delta G_{exp}\rvert$ is the absolute difference between the calculated and experimental values.

molecular dynamics simulation, which is performed for the wild-type system. This typically takes about 80−90% of the total time required to evaluate by computational ASM a typically sized protein−protein interface. Only the remaining 10−20% of the computational time grows linearly with the number of mutations evaluated. Hence, considering 10 or 100 mutations in ASM does not significantly increase the CPU time associated, whereas in TI would imply a 10-fold increase.

With this ASM protocol, we can easily and quickly calculate the difference in free energy of binding, upon mutation of several interfacial amino acid residues from a fast dynamics in implicit solvent and using the VMD plugin within the same time frame that is required in TI to evaluate a single mutation.

IV. CONCLUSIONS

In this study, we have calculated the protein−protein binding free energy differences upon alanine mutation of interfacial residues ($\Delta\Delta G_{bind}$) both with the computational ASM protocol and TI, for 22 critical mutations for which accurate experimental data was available.

Even though the present test set can be regarded as relatively small, it involves quite diverse mutations, representative in terms of type and range of energy value associated to those typically encountered when studying protein−protein interfaces by experimental ASM. Hence, we feel confident about the conclusion derived from this comparison.

Globally, the results show that this faster and easier computational ASM protocol is capable of reproducing experimental mutagenesis results with good accuracy, at the same level of accuracy of TI, and its use is very appealing in the systematic study of protein−protein interfaces. Naturally, TI has a wider range of applications in the sense that it can be applied in the study of other mutations (not only alanine scanning mutagenesis) and in more general applications, with a high computational cost associated.

■ ASSOCIATED CONTENT

Supporting Information

VEGF/FLT-1 interface (Figure S1), RMSd analysis of the VEGF/FLT-1 complex in the MD simulations (explicit and implicit solvent) (Figure S2) and free energies of each substep obtained with TI (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pafernan@fc.up.pt.

Author Contributions

†These authors have contributed equally to this paper

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are thankful for the financial support provided by FCT (PTDC/QUI-QUI/100372/2008, PTDC/QUI-QUI/102760/2008, SFRH/BD/46867/2008, SFRH/BD/43600/2008, and Grant No. Pest-C/EQB/LA0006/2011).

REFERENCES

(1) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot spots—A review of the protein−protein interface determinant amino-acid residues. *Proteins* **2007**, *68*, 803−812.

(2) Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, *280*, 1−9.

(3) Higueruelo, A. P.; Schreyer, A.; Bickerton, G. R. J.; Pitt, W. R.; Groom, C. R.; Blundell, T. L. Atomic interactions and profile of small molecules disrupting protein−protein interfaces: The TIMBAL Database. *Chem. Biol. Drug Des.* **2009**, *74*, 457−467.

(4) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Computational alanine scanning mutagenesis—An improved methodological approach. *J. Comput. Chem.* **2007**, *28*, 644−654.

(5) Tuncbag, N.; Gursoy, A.; Keskin, O. Identification of computational hot spots in protein interfaces: Combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* **2009**, *25*, 1513−1520.

(6) Xia, J. F.; Zhao, X. M.; Song, J. N.; Huang, D. S. APIS: Accurate prediction of hot spots in protein interfaces by combining protrusion index with solvent accessibility. *BMC Bioinf.* **2010**, *11*, 174.

(7) Darnell, S. J.; LeGault, L.; Mitchell, J. C. KFC Server: Interactive forecasting of protein interaction hot spots. *Nucleic Acids Res.* **2008**, *36*, W265−W269.

(8) Darnell, S. J.; Page, D.; Mitchell, J. C. An automated decision-tree approach to predicting protein interaction hot spots. *Proteins* **2007**, *68*, 813−823.

(9) Cho, K.-i.; Kim, D.; Lee, D. A feature-based approach to modeling protein−protein interaction hot spots. *Nucleic Acids Res.* **2009**, *37*, 2672−2687.

(10) Liu, Q. A.; Li, J. Y. Protein binding hot spots and the residue-residue pairing preference: a water exclusion perspective. *BMC Bioinf.* *11*, 244.

(11) Cho, K. I.; Kim, D.; Lee, D. A feature-based approach to modeling protein−protein interaction hot spots. *Nucleic Acids Res.* **2009**, *37*, 2672−2687.

(12) Guharoy, M.; Chakrabarti, P. Empirical estimation of the energetic contribution of individual interface residues in structures of protein−protein complexes. *J. Comput.−Aided Mol. Des.* **2009**, *23*, 645−654.

(13) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.

(14) Massova, I.; Kollman, P. A. Computational alanine scanning to probe protein−protein interactions: A novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.

(15) Huo, S.; Massova, I.; Kollman, P. A. Computational alanine scanning of the 1:1 human growth hormone−receptor complex. *J. Comput. Chem.* **2002**, *23*, 15−27.

(16) Wiesmann, C.; Fuh, G.; Christinger, H. W.; Eigenbrot, C.; Wells, J. A.; de Vos, A. M. Crystal structure at 1.7 Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell* **1997**, *91*, 695−704.

(17) Buckle, A. M.; Schreiber, G.; Fersht, A. R. Protein−protein recognition—Crystal structural-analysis of a Barnase Barstar complex at 2.0 Å resolution. *Biochemistry* **1994**, *33*, 8878−8889.

(18) Mosyak, L.; Zhang, Y.; Glasfeld, E.; Haney, S.; Stahl, M.; Seehra, J.; Somers, W. S. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J.* **2000**, *19*, 3179−3191.

(19) Bhat, T. N.; Bentley, G. A.; Boulot, G.; Greene, M. I.; Tello, D.; Dallacqua, W.; Souchon, H.; Schwarz, F. P.; Mariuzza, R. A.; Poljak, R. J. Bound water molecules and conformational stabilization help mediate an antigen−antibody association. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1089−1093.

(20) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An automated pipeline for the setup of Poisson−Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665−W667.

(21) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very fast prediction and rationalization of pKa values for protein−ligand complexes. *Proteins* **2008**, *73*, 765−783.

(22) Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and rationalization of protein pK_a values. *Proteins* **2005**, *61*, 704−721.

(23) Olsson, M. H. M.; Sǎndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.

(24) Case, D. A.; Darden, T. Cheatham, T. E., III; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K.; Paesani, F. Wu, X. Brozell, S. Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, 2006.

(25) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(26) Tsui, V.; Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **2001**, *56*, 275−291.

(27) Izaguirre, J. A. Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **2001**, *114*, 2090−2098.

(28) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin dynamics of peptides—The frictional dependence of isomerization rates of *n*-acetylalanyl-*n*-methylamide. *Biopolymers* **1992**, *32*, 523−535.

(29) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald—An *n* log(*n*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(30) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of Cartesian equations of motion of a system with constraints—Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(31) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Detailed microscopic study of the full ZipA: FtsZ interface. *Proteins* **2006**, *63*, 811−821.

(32) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Unraveling the importance of protein−protein interaction: Application of a computational alanine-scanning mutagenesis to the study of the IgG1 streptococcal protein G (C2 fragment) complex. *J. Phys. Chem. B* **2006**, *110*, 10962−10969.

(33) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Unravelling hot spots: A comprehensive computational mutagenesis study. *Theor. Chem. Acc.* **2007**, *117*, 99−113.

(34) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot spot computational identification: Application to the complex formed between the hen egg white lysozyme (HEL) and the antibody HyHEL-10. *Int. J. Quantum Chem.* **2007**, *107*, 299−310.

(35) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Backbone importance for protein−protein binding. *J. Chem. Theory Comput.* **2007**, *3*, 885−893.

(36) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot spot occlusion from bulk water: A comprehensive study of the complex between the lysozyme HEL and the antibody FVD1.3. *J. Phys. Chem. B* **2007**, *111*, 2697−2706.

(37) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Protein−protein recognition: A computational mutagenesis study of the MDM2-P53 complex. *Theor. Chem. Acc.* **2008**, *120*, 533−542.

(38) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330−14335.

(39) Bradshaw, R. T.; Patel, B. H.; Tate, E. W.; Leatherbarrow, R. J.; Gould, I. R. Comparing experimental and computational alanine

scanning techniques for probing a prototypical protein−protein interaction. *Protein Eng. Des. Sel.* **2011**, *24*, 197−207.
(40) Rocchia, W.; Alexov, E.; Honig, B. Extending the applicability of the nonlinear Poisson−Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B* **2001**, *105*, 6507−6514.
(41) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *J. Comput. Chem.* **2002**, *23*, 128−137.
(42) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Accuracy of the numerical solution of the Poisson−Boltzmann equation. *J. Mol. Struc.—Theochem* **2005**, *729*, 11−18.
(43) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(44) Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **1983**, *16*, 548−558.
(45) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics Modell.* **1996**, *14*, 33−38.
(46) Ribeiro, J. V.; Cerqueira, N. M. F. S. A.; Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. CompASM: an Amber-VMD alanine scanning mutagenesis plug-in. *Theor. Chem. Acc.* **2012**, *131*, 1271.
(47) Case, D. A.; Darden, T.; Cheatham, T. E., III; Simmerling, C. Wang, J. Duke, R. E.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W. Merz, K. M..Wang, B. Hayik, S. Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V. Cui, G. Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, 2008.
(48) Hawkins, R. E.; Russell, S. J.; Baier, M.; Winter, G. The contribution of contact and noncontact residues of antibody in the affinity of binding to antigen—The interaction of mutant D1.3 antibodies with lysozyme. *J. Mol. Biol.* **1993**, *234*, 958−964.
(49) DallAcqua, W.; Goldman, E. R.; Eisenstein, E.; Mariuzza, R. A. A mutational analysis of the binding of two different proteins to the same antibody. *Biochemistry* **1996**, *35*, 9667−9676.
(50) Schreiber, G.; Fersht, A. R. Interaction of Barnase with its polypeptide inhibitor Barstar studied by protein engineering. *Biochemistry* **1993**, *32*, 5145−5150.
(51) Keyt, B. A.; Nguyen, H. V.; Berleau, L. T.; Duarte, C. M.; Park, J.; Chen, H.; Ferrara, N. Identification of vascular endothelial growth factor determinants for binding KDR and FLT-1 receptors—Generation of receptor-selective VEGF variants by site-directed mutagenesis. *J. Biol. Chem.* **1996**, *271*, 5638−5646.
(52) Schreiber, G.; Fersht, A. R. Energetics of protein−protein interactions—Analysis of the Barnase−Barstar interface by single mutations and double mutant cycles. *J. Mol. Biol.* **1995**, *248*, 478−486.
(53) Perez, M. A.; Sousa, S. F.; Oliveira, E. F.; Fernandes, P. A.; Ramos, M. J. Detection of farnesyltransferase interface hot spots through computational alanine scanning mutagenesis. *J. Phys. Chem. B* **2011**, *115*, 15339−15354.