Effect of sampling on BACE-1 ligands binding free energy predictions via MM-PBSA calculations

Nicolas Chéron, Eugene I Shakhnovich

Effect of sampling on BACE-1 ligands binding free energy predictions *via* MM-PBSA calculations

[1,2]Nicolas Chéron and [1]Eugene I. Shakhnovich**

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge MA 02138, USA; [2]Ecole Normale Supérieure, PSL Research University, UPMC Univ. Paris 06, CNRS, Département de Chimie, UMR 8640 PASTEUR, 24 rue Lhomond, 75005 Paris, France.

**Correspondence: nicolas.cheron@ens.fr ; shakhnovich@chemistry.harvard.edu.*

Abstract: The BACE-1 enzyme is a prime target to find a cure to Alzheimer's disease. In this paper, we used the MM-PBSA approach to compute the binding free energies of 46 reported ligands to this enzyme. After showing that the most probable protonation state of the catalytic dyad is mono-protonated (on ASP32), we performed a thorough analysis of the parameters influencing the sampling of the conformational space (in total, more than 35 μs of simulations were performed). We show that ten simulations of 2ns gives better results than one of 50ns. We also investigated the influence of the protein force field, the water model, the periodic boundary conditions artifacts (box size), as well as the ionic strength. Amber03 with TIP3P, a minimal distance of 1.0nm between the protein and the box edges and a ionic strength of I=0.2M provides the optimal correlation with experiments. Overall, when using these parameters, a Pearson correlation coefficient of R=0.84 (R^2=0.71) is obtained for the 46 ligands, spanning 8 orders of magnitude of K_d (from 0.017nm to 2000μM i.e. from -14.7 to -3.7 kcal/mol), with a ligand size from 22 to 136 atoms (from 138 to 937 g/mol). After a two-parameter fit of the binding affinities for 12 of the ligands, an error of RMSD = 1.7 kcal/mol was obtained for the remaining ligands.

Keywords: Drug design, protein-ligand interaction, molecular dynamics, free energy, MM-PBSA

Introduction

According to the Alzheimer's Disease International association, the global cost of dementia in 2015 is US $818 billion, making it one of the costliest disease. The most common cause of dementia is Alzheimer's disease, for which there is still no treatment. Two hypotheses have been formulated to explain the cause of this disease: the formation of amyloid β-peptide plaques or the lack of stabilization of microtubules due to the hyperphosphorylation of tau proteins (the two being not mutually exclusive)[1,2]. The formation of the amyloid β-peptides is due to one specific enzyme, the β-site Amyloid precursor protein Cleaving Enzyme 1, also called BACE-1 or β-secretase[3,4]. Significant efforts have been focused towards the inhibition of this enzyme, aiming at finding a cure to Alzheimer's disease[5,6]. Monoclonal antibodies have also been investigated for this disease, but the recent failure of the solanezumab program from Eli Lilly raises questions about this approach. To date no drug has been approved for the BACE-1 target protein; nonetheless, several programs are in phase 2/3 clinical trials, such as one from Merck (for verubecestat or MK-8931)[7], one from AstraZeneca and Eli Lilly (AZD3293) or one from Johnson & Johnson (JNJ-54861911). From a pragmatic point of view, one must hope that these molecules will be efficient cures to Alzheimer's disease and will be approved. However, it does not mean that the quest for BACE-1 inhibitors should stop: firstly, these programs may fail and thus new potent molecules are needed; secondly, more drugs on the market means competitions amongst them and thus lower costs for both the patient and the healthcare system; thirdly, approved drugs may be not efficient for sub-classes of patients bearing for example a mutation of the target protein. Thus, there is still a need to discover new molecules capable of inhibiting the BACE-1 enzyme.

The road to discover a new drug is long and full of obstacles; it usually starts from the identification of a first hit that is further optimized to a lead, and many properties (pharmacokinetic

and pharmacodynamic) must be tuned at the same time. Even if some compromises must be made, there is one central property that must be always kept as high as possible: the affinity of the ligand to the target protein (expressed as binding free energy or dissociation constant, with $\Delta G_{Binding}=RT*\ln(K_d)$). In a drug discovery program where many compounds (from hundreds to millions) are considered, computational tools can play a crucial role in sorting molecules and help identifying promising candidates to circumvent costly experiments. For example, tens of millions of compounds can now be docked within a month. However, results from docking calculations need to be refined since even if the orientation of ligands in the active site can now be predicted efficiently, the accuracy of scoring functions is limited[8]. In the present work, our goal has been to find a fast yet accurate approach to estimate the binding free energy of ligands with the human BACE-1 enzyme that would be valid in a wide range of values (of $\Delta G_{Binding}$) and for a diverse set of ligands, in order to use it as an extra filtering tool for sorting molecules before final synthesis and testing. We used protein-ligand crystallographic structures as starting points in order to benchmark the approach, that can later be used with docked structures.

The most accurate methods to calculate binding free energy rely on all-atom molecular simulations (molecular dynamics (MD) or Monte-Carlo), which can provide statistically meaningful conformational ensembles. Scoring functions can be used to estimate $\Delta G_{Binding}$ on a large scale, but their predictive power is too low when it comes to selecting molecules for experimental assays (correlation between scores and experiments of at most $R^2=0.38$ currently[8]). There are roughly two families of approaches that use simulations to compute binding free energies: alchemical transformations (such as Free Energy Perturbation (FEP) or Thermodynamic Integration (TI)) and end-point methods (such as Molecular Mechanics Poisson Boltzmann Surface Area or MM-PBSA)[9]. FEP or TI are exact in theory (since they are derived from exact

statistical mechanical relationships) and are usually deemed to be more accurate. However, it comes at the cost of excessive simulation time since they can be hard to converge, especially when applied to a set of diverse ligands. Thus, they cannot be used routinely to sort on the order of 1000 unknown ligands. FEP was recently applied to the BACE-1 protein on a family of 32 related ligands, leading to a correlation of R^2=0.69 between experiments and computations[10]. Alternative approaches based on MM-GBSA and MM-PBSA have been extensively documented and used[9,11–15], and we will not review them in detail here (a summary is provided in Supporting Information). In these methods, two criteria are of importance: the sampling and the energy calculations. There have been several reports that focus on the latter (i.e. the MM-PBSA parameters, see [16–18] e.g.), but fewer that target the former (the MD protocols[19]). Thus, we report in this paper the search for an optimal set of simulation protocols to sample the conformational space of the BACE-1 enzyme as accurately as possible using MD simulations; MM-PBSA calculations are then used to estimate $\Delta G_{Binding}$. In the present paper, we didn't investigate the MM-GBSA method based on the Generalized Born (GB) model; MM-GBSA has been widely used since it is faster than MM-PBSA, but the gain of speed of computers has minimized the benefits of that gain. We have investigated the effect of the computational protocols including simulation length, the protein force field, the water model, the box size and the ionic strength, on the Pearson correlation coefficient between predicted and experimental values of the binding affinity.

Preparation of the ligands

We started by looking for the keywords "Human Beta Secretase 1", "BACE-1", "secretase beta" and "Beta-secretase 1" in the PDBBind-CN database[20]. Out of the 264 results, 54 complexes were found with K_d data. After removing those with a resolution worse than 2.5 Å or for which we couldn't verify the data, we kept 46 structures. From these crystallographic structures, we prepared

an initial set of 12 complexes (used to find the parameters that lead to the highest correlation between computational and experimental data) that were selected amongst those with no or only a few missing atoms in the crystallographic structure, targeting a large range of K_d (from 0.017nM to 200µM i.e. from -14.7 to -5.0 kcal/mol, obtained with $\Delta G=RT*\ln(K_d)$ with R the gas constant and T=298.15K) and of ligand size (from 29 to 136 atoms and from 233 to 937 g/mol). They are listed in Table 1. Once the optimal parameters were found, we computed the binding energy for the remaining 34 complexes (also listed in Table 1). Since crystallographic structures can have missing atoms or residues, all complexes were pre-processed with UCSF Chimera[21]. The interface of Modeller was used[22] (Model/Refine Loops) with default options (model « non-terminal missing structures », generate 5 models, standard « loop modeling protocol »). We always kept the one with the lowest zDOPE score. Missing atoms and ligand hydrogen atoms were then added with the Dock Prep tool of Chimera[21], keeping default options (don't delete solvent, keep highest occupancy, incomplete side chains with Dunbrack rotamers library[23], add hydrogens, don't add charges). When the crystallographic structure was a dimer, trimer or tetramer, we kept only one chain (the one with the fewest missing atoms); we also removed molecules different than the protein, the ligand and water (such as tartric acid, ions, e.g.). 40 of the 46 proteins share the same amino acids sequence (they are different only by the sequence length). The ID 3bra, 3buf, 3bug and 3buh are a single mutant (K252A) whereas 2q11 and 2q15 are a quadruple mutant (N98Q/N117Q/N168Q/N299Q). Ligand's geometries were extracted from the crystallographic structures and hydrogen atoms were added using UCSF Chimera[21]. The structures were then manually checked to assign a protonation state corresponding to pH=7. Chemical structures of the ligands are displayed in Supporting Information.

Computational details

MD simulations were performed on GPU cards with GROMACS 5.1.2[48–51]. Ligand's geometry were first optimized at the M06-2X/6-31+G** level of theory with an implicit description of water (IEF-PCM) before calculating the RESP charges at the HF/6-31G* level of theory with the Merz-Kollman scheme (with the Gaussian 09 software[52]), consistently with the approach used by the general AMBER force field (GAFF)[53]. Acpype[54] (an interface to antechamber[55]) was then used to generate the ligand topology using GAFF[53]. Our standard protocol consisted of solvating the system in a rhombic dodecahedron box with a minimal distance of 1.0 nm between the complex and the box edges (unless otherwise noted), following the parameters used by Amber03[56] (whereas during the development of Amber14SB, a minimal distance of 0.8 nm was used). Minimal number of ions (chloride or sodium) was then added to neutralize the box; when we investigated the role of ionic strength, we also added ions at a given concentration in addition to the minimal amount needed. The system energy was then minimized before performing a primary NPT simulation of 500ps (δt=1fs, velocity-rescale thermostat[57], Berendsen barostat[58]) during which the temperature increased from 100 to 300K in 400ps *via* simulated annealing. We then performed our production simulation in the NPT ensemble at 300K (δt=2fs, velocity-rescale thermostat, Parinello-Rahman barostat[59]). The first ns of the production simulation was always excluded from the analysis to allow for further equilibration. A pseudo-random number seed was used for both the stochastic term in the velocity-rescale thermostat and the initialization of velocities. Bonds containing a hydrogen atom were constrained during the simulations using the LINCS algorithm[60,61]. By default, the Amber03 force field[56] was used together with the TIP3P water model[62,63]. We also investigated the role of Amber99SB[64] and Amber14SB[65] protein force fields, as well as SPC/E[66], TIP4P[63], TIP4P-2005[67] and TIP5P[68] water models. For the three protein force fields, we used the standard parameters regarding non-

bonding interactions, i.e. an 8 Å cutoff for van der Waals interactions and a PME[69] description of electrostatic ones with a similar value of 8 Å for the separation between direct and reciprocal spaces.

Unless otherwise mentioned, we used the so-called "single trajectory approximation" to compute the binding free energies where structures for the ligand, the protein and the complex are coming from the same trajectory. Snapshots of the protein-ligand complex were extracted every 10ps (i.e. 100 per ns), always leaving the first ns of the simulation aside for further equilibration, and the AMBER12[70] package was used to compute the MM-PBSA binding energy[71] (except for simulations with Amber14SB where AmberTools16 was used for the MM-PBSA calculations). Defaults parameters were used for the probe size (1.4 Å) and the radii (mbondi set). The surface tension and offset of the cavity of the Poisson−Boltzmann calculations were set to 0.0072 kcal/mol/Å^2 and 0.00 kcal/mol respectively[13,72,73], and the nonpolar optimization method of AMBER12 ("inp" keyword) was set to 1 meaning that "the total non-polar solvation free energy is modeled as a single term linearly proportional to the solvent accessible surface area" (Amber12 manual). As previously reported, the use of the default value of 2 (where "the total non-polar solvation free energy is modeled as two terms: the cavity term and the dispersion term") does not always allow to reproduce experimental results[74] or leads to lower correlations[18]. By default, the ionic strength for Poisson−Boltzmann calculations was set to 0.0M (unless otherwise mentioned) to match the ionic strength used in MD simulations. When reporting MM-PBSA values, an arithmetic average is always used for both the binding term (MM) and the solvation term (PBSA). We discuss in the Supporting Information the influence of using other types of averages. Raw MM-PBSA data with uncertainties are provided in Supporting Information. To complete MM-PBSA calculations, an entropy contribution is sometimes evaluated to consider translational,

rotational and vibrational terms. However, a preliminary study performed on seven complexes, using a normal mode analysis on snapshots extracted every 50ps (i.e. 20 per ns) revealed that the entropy contribution does not affect the correlation between computation and experimental values in the present case. Indeed, we found a correlation of R^2=0.79 without the entropy and R^2=0.76 with the entropy term. In other cases previously studied (such as the HIV-1 protease), we found that the correlations are respectively R^2=0.48 without the entropy and R^2=0.68 with it (see [74]); thus, including or not an entropy term is system-dependent and we note that it is still difficult to know in advance if there is a gain in including this term or not. Moreover, entropic values can fluctuate a lot and some authors prefer not to include them[12]. Considering that entropy calculations can be time-consuming (especially when the system is large as the BACE-1 protein, since it scales as $(3N)^3$ where N is the number of atoms), that our goal here was to find a fast yet accurate method and that using the entropy term does not improve the correlation, we have decided not to include them in this study.

Errors between the computed MM-PBSA values and experiments were estimated by the root-mean-square deviation (RMSD) and the standard deviation (SD), which respectively represent the deviations to experiments and to a linear fit, and are calculated as:

$$RMSD = \sqrt{\frac{\sum\left[\Delta G_{MM-PBSA} - \Delta G_{Experiments}\right]^2}{N}}$$

$$SD = \sqrt{\frac{\sum\left[\Delta G_{MM-PBSA} - \left(a + b * \Delta G_{Experiments}\right)\right]^2}{N-1}}$$

where *a* and *b* are the intercept and slope of the linear regression. To obtain absolute binding free energies with MM-PBSA calculations, one needs to correct the results with terms for standard concentration or pressure and volume to obtain absolute energetic values[9,18], in addition to the

entropic terms. Thus, the MM-PBSA approach used here cannot lead to absolute values but only to relative ones and the reported RMSD can seem to be high. However, comparisons between RMSD for different methods are still valid. A "corrected RMSD" (noted cRMSD), where for each set of parameters the systematic error (i.e. the mean signed error) is subtracted from the MM-PBSA values before computing the RMSD, is also provided. This allows to obtain an estimate of the error in relative binding free energies.

$$cRMSD = \sqrt{\frac{\sum\left[\left(\Delta G_{MM-PBSA} - \Delta G_{Experiments}\right) - \left(\langle \Delta G_{MM-PBSA} \rangle - \langle \Delta G_{Experiments} \rangle\right)\right]^2}{N}}$$

Similarly, the value of the standard deviation is highly dependent on the slope of the linear regression fit. Thus, we also provide "corrected SD" values (noted cSD) where the MM-PBSA values are scaled to place them in the same range as the experimental values (where $\langle ... \rangle$ represents averages over the set).

$$\Delta G'_{MM-PBSA} = \frac{\langle \Delta G_{Experiments} \rangle}{\langle \Delta G_{MM-PBSA} \rangle} \Delta G_{MM-PBSA}$$

$$cSD = \sqrt{\frac{\sum\left[\Delta G'_{MM-PBSA} - \left(a' + b' * \Delta G_{Experiments}\right)\right]^2}{N-1}}$$

where a' and b' are the intercept and slope of the linear regression between $\Delta G'_{MM-PBSA}$ and ΔG_{Exp}.

Uncertainties of the correlation coefficient were calculated using a bootstrapping analysis on all the independent simulations with 100 cycles. In all cases, the statistical error on R^2 (calculated as the Student's 95% confidence interval) was always lower than 0.002 and thus is not reported. The correlation between experiments and computations also depends on uncertainties in the experimental values. Unfortunately, this value is not available for all complexes. To take it into account, we performed a bootstrapping analysis where each experimental value is randomly modified by a random number coming from a normal distribution with a standard deviation of 0.5

kcal/mol (assuming a standard uncertainty in experiments of 0.5 kcal/mol). In all cases, the statistical error on R^2 was lower than 0.008. The statistical errors computed for cRMSD and cSD (with the bootstrapping analysis where experimental data are modified by random numbers) are respectively always lower than 0.06 and 0.05 kcal/mol. As a consequence, differences of 0.01 unit of R^2 are statistically significant, as well as differences of 0.1 kcal/mol for cRMSD and cSD. In the following, unmodified experimental results are used.

Protonation state of the active site and simulation length

The active site of the BACE-1 protein contains two aspartate residues (ASP32 and ASP228) and each one can be either protonated or unprotonated, leading to four possible states (see Figure 1). Almost all options have been reported in the literature: a monoprotonated dyad (where only ASP32[75–77] –which tends to become the consensus— or ASP228[78] is protonated), a di-protonated dyad[79], as well as the possibility that the protonation depends on the ligand[80,81]. One must note that most of the studies leading to these conclusions relied on structural (*via* MD simulations) and docking procedures.

To assess the correct protonation states, we investigated which state leads to the highest correlation between computational and experimental values: for each complex in the initial set, we performed simulations with the four possible states. For each state, we performed 50 independent simulations of 5ns (that are different only by the initial velocities that are randomly assigned) and calculated the averaged MM-PBSA energies. We verified that the distributions of all the MM-PBSA energies (from 5000 to 20000 for each complex and each state) follow a normal distribution (see Supporting Information). All the correlations are reported on Table 2, and scatter plots are available in Supporting Information. We took the opportunity to also investigate the needed simulation length to obtain converged results by calculating the MM-PBSA energies on different blocks of the simulations. First, one can note that a simulation length of 2ns is enough since for

each state, all correlations are roughly the same whatever the time block used. Then, a significant difference is found between the states: the highest correlation and lowest errors between computation and experiments values (R=0.93, R^2=0.86) is found with State2 (where only ASP32 is protonated) or when the lowest binding energy from the four states is kept ("BestState") which means that the protonation state could depend on the ligand. However, correlation coefficients are slightly higher with State2 than with BestState (we recall here that statistical error is always below 0.002 units of R^2) and RMSD is lower by 2 kcal/mol for State2 (whereas the differences in the other computed errors are below 0.2 kcal/mol). Since using a single protonation state is more straightforward to apply (and more easily generalizable) to a full set of 46 ligands, we used the State2 for the remaining of the study, in agreement with previous results[75–77].

We have used a procedure with 50 independent replicas because it has been shown that it is a good procedure to obtain proper sampling[82] since it avoids being trapped in a conformational well. We compared this approach with one using a single trajectory of 51ns, using the last 50ns to compute the MM-PBSA energies and the same number of snapshots: using State2 for the protonation with the same parameters as in Table 2, we obtained a correlation of R^2=0.79, and errors of RMSD=36.3 kcal/mol, cRMSD=18.6 kcal/mol, SD=10.1 kcal/mol and cSD=2.5 kcal/mol (we recall that MM-PBSA provides here relative values which is the reason why we propose the use of cRMSD and cSD to obtain more reliable estimates of errors). Thus, the correlation is slightly lower than with multiple replicas with significantly higher errors. As a consequence, we have decided to keep using an approach with several short simulations instead of a unique long since it provides overall better results (similar conclusions were obtained before[17,83], and it is likely that this statement is generally applicable).

Number of replicas

Since a better correlation is achieved when performing many short simulations instead of one very long, we studied the number of simulations needed to converge results by using a bootstrapping approach after performing 100 independent simulations of 2ns. We computed the average binding energies over N randomly selected simulations over the set of 100 for each complex, and calculated the correlation with experiments. 100 cycles of bootstrapping were performed for each size, which allowed us to compute average correlation coefficients with the associated errors. For N from 10 to 100, $<R^2>$ is almost constant since it goes from 0.85 to 0.86 (with an error determined as the Student's 95% confidence interval going from 0.004 to 0.001, see Table 3). Thus, when using 10 replicas of 2ns the correlation is already converged (and gives better correlation than with one simulation of 51ns); however, to minimize the statistical error in the current study, we always used 50 simulations in the following since the error significantly decreases with more replicas.

Protein force field

After having found the optimal simulation time and number of replicas, we started to compare parameters used during the simulations for the sampling of the conformational space. Firstly, three protein force fields were considered: Amber03, Amber99SB and Amber14SB (see Table 4). It appears that the best correlation is obtained with the Amber03 force field (from 2003), which is surprising since it is older than Amber99SB (2006) and Amber14SB (2014) and we usually expect more recent force fields to be more accurate. Moreover, the errors are always lower for the Amber03 force field. However, the differences between Amber03 and Amber99SB are small and may also come from the experimental incertitude. Other studies also reported low differences when using different force fields[14,19], and sometimes better results of MM-PBSA with older force fields (Amber03 vs Amber99SB, e.g.). We concur with the analysis of Xu *et al.* when

they state that "the fitting for new force field parameters primarily emphasizes the conformations and dynamics of proteins, not the binding free energies"[14], which explain why older force fields may be more suited for MM-PBSA calculations. For the remaining part of the study, we used Amber03.

Water model

Protein force fields are developed with a given choice of water model, TIP3P for the Amber family of force fields. However, the TIP3P model is known to have several caveats, such as wrong structures and dynamics[84] and a wrong phase diagram[85]. Thus, the choice of a water model is a matter of compromise between being consistent with the protein force field and having a correct description of the water itself. We investigated the effect of five different water models on the prediction of experimental results (see Table 5). The two three points models (TIP3P and SPC/E) give very similar results, both in terms of correlation coefficient than for errors. The four points models provide a lower correlation, but also lower RMSDs and higher SDs. TIP5P gives a correlation similar to TIP3P and SPC/E with slightly higher errors and at the cost of worse performances. These results show it is more important to use the same protocol as the one used during the development of force fields, rather than improving the description of the solvent and its dynamics with better models, which is a conclusion that can be applied to other systems. Thus, we have decided to keep the TIP3P water model for the remaining of the study and note that SPC/E would be as valid as TIP3P.

Box sizes and cut-off influences

Force fields are parameterized with a given size of box and cutoff values for the non-bonded interactions. However, to minimize artifacts of periodic boundary conditions used in MD simulations, one may want to increase the box size or the cutoff values. As for the water model,

the choice is then a matter of compromise and we wanted to investigate here the consequences of this choice. We report in Table 6 results for three sizes of box that we call small (used thus far), medium and large. It appears that results are independent of the box size and cutoff values, and trying to reduce the artefacts of periodic boundary conditions does not improve the correlation. Thus, for box size and simulation cutoffs, we kept the parameters used during the development of the force field (small box) and this result is probably transferable to other systems.

Ionic strength

Another parameter of importance during the simulations is the number of ions included in the system. In our protocol, there are two ionic strengths to consider: the one used during the molecular dynamics simulations and the one used for Poisson-Boltzmann calculations. The standard approach in biomolecular simulations is usually to only add the minimal number of ions needed to neutralize the box. However, this is not a good modelling of experimental setups. Indeed, a buffer is very often used experimentally when measuring the experimental binding free energies. For example, for the 4fs4 complex, the buffer consisted of "100 mM sodium acetate pH 3.5, 100 mM NaCl, 0.0125% Brij-35, supplemented with 1.5% DMSO"[34]. Unfortunately, these experimental details are not always reported. The presence of ions will change the ionic strength, and may have an influence on the conformational space visited by the system during the simulations. Since a proper force field may not be available for all the molecules contained in the buffers, a solution is to model the buffer with a sodium chloride solution of the same ionic strength. However, even when experimental details are provided it can be impossible to convert them to an ionic strength. Thus, in complement to the previous simulations where no extra ions were added to the system (except for neutralization) and where the ionic strength for MM-PBSA was set to I=0M, we have performed simulations where a concentration of sodium chloride of 0.1M, 0.2M

and 0.3M was added. Our goal was to find an ionic strength that leads to the best modelling of experimental setups. For each simulation, we performed the MM-PBSA calculations with an ionic strength of 0.0, 0.1, 0.2 and 0.3M to separate the influence of the ions in the sampling and in the estimation of the binding energy. In Table 7, we report the correlation (R^2) for all the combinations, and we can see that the ionic strength used in the MM-PBSA calculations has no influence on the estimated binding energies. However, the amount of ions used during the simulations has a very little influence on it (from 0.85 with [NaCl]=0.3M to 0.88 with [NaCl]=0.2M). Thus, it is most appropriate to use a concentration of [NaCl]=0.2M during the MD simulations and I=0.2M for the MM-PBSA calculations, since it is closest to experimental conditions and allows to obtain the highest correlation, as well as lowest errors.

Single trajectory approximation

To study a binding process, one needs to make a difference between the energy of the complex and the energy of the protein and the ligand separated. In most of the work with MM-PBSA, a single trajectory is used: a simulation of the protein-ligand complex is performed, and conformations of the protein and ligand are extracted from the complex conformations. There are two main reasons for using this approximation: (1) speed and simplicity: only one simulation is performed. However, the simulation time for the ligand isolated is usually negligible with regards to the time for the complex or the protein. Moreover, when performing virtual screening of ligands against one protein, the protein is always the same. Thus only one extra simulation for the protein is needed and the additional cost of using multiple trajectories is low. When the protein is known to undergo conformational change upon binding, it is usually considered mandatory to use multiple trajectories. However, if only relative free energies are needed for the binding to the same protein, the contribution of the apo-protein will be the same for all the ligands and one does not need to

perform the extra simulation. This of course does not apply when absolute energies are sought, when mutants of the protein are considered, or when all the crystallized structures do not have the same sequence length, as is the case here. (2) Numerical noise: the potential terms for internal energy (bonds, angles, dihedral angles) are noisy. When a single trajectory is used, they all cancel each other, leading to less fluctuation, and better results are often obtained with a single trajectory. To minimize the problem of numerical noise with multiple trajectories, longer simulation times are needed to obtain a better convergence.

We assessed whether better correlations could be achieved without using the single trajectory approximation for the BACE-1 enzyme. We performed 50 simulations of the protein alone and of the ligand alone; the complex simulations were 5ns long (using the previous data from the last 4ns), and the protein and ligand simulations were all 10ns (using data from the last 5ns). Since for each complex the protein sequence lengths are different, we performed simulations for all the isolated proteins. We present in Table 8 the correlations and errors found in different cases: the single trajectory approximation, the use of multiple (three) independent trajectories, as well as hybrid cases where ligand and complex conformations are coming from the same trajectory and protein conformations are sampled separately, and one case where protein and complex are coming from the same trajectory and ligand conformations are sampled separately. Surprisingly, the use of multiple trajectories leads to a very low correlation between computational and experimental results (R^2=0.11, with a statistical error estimated as the Student's 95% confidence interval of 0.02). When compared with the hybrid cases, this appears to be because of the different conformational space adopted by the protein since when the ligand (and only the ligand) energies are coming from independent simulations the correlation is similar to the single trajectory (0.83 vs 0.86) with larger errors though. The reason for the poor performance of the multiple trajectory may

be that the apo-protein can adopt more conformations than the complex, and even longer simulations would be needed. However, since we were looking for a fast approach, we consider that in the present case of the BACE-1 enzyme, the use of the single trajectory approximation is more suited.

Full set of ligands

Finally, we computed the energy of binding for the remaining 34 complexes (see Figure 2). 50 simulations of 2ns with the protonation state State2 (only ASP32 is protonated) was used, with the Amber03 force field for the protein and TIP3P for the water model, a minimal distance of 1.0 nm between the protein and box edges, [NaCl]=0.2M during MD simulations and I=0.2M for MM-PBSA calculations. As expected with a larger dataset the overall correlation decreases, but is still very high: R=0.84 (R^2=0.71). A bootstrapping analysis similar to the one presented on Table 3 (on the 50 values with 100 cycles) was performed and provided error bars for the correlation: 0.0005 for R and 0.0009 for R^2. Calculated errors are RMSD=26.7 kcal/mol, cRMSD=12.9 kcal/mol, SD=8.3 kcal/mol and cSD=2.3 kcal/mol. When the MM-PBSA energies are calculated on only the minimized crystallographic structure, the correlation with experiments is R^2=0.66, lower than with the MD procotol. Moreover, errors are higher: RMSD=30.4 kcal/mol, cRMSD=11.3 kcal/mol, SD=24.7 kcal/mol and cSD=3.5 kcal/mol. Thus, if one wants to compute the energy of a ligand in its crystal structure conformation it is possible to use only energy minimization (with large errors though). However, when new ligands with unknown crystal structure are sought and virtual screening is used, it is always necessary to refine the docked poses to assess their stability, for example by performing simulations. Thus, for putative new ligands of the BACE-1 protein for which no crystallographic structures are known, molecular dynamics

simulations can be used after docking and the workflow presented here can be followed to compute the binding free energy.

Another approach can be followed to illustrate that this protocol can be used to predict binding free energies. Using only the initial set of 12 ligands with the same parameters as the ones from Figure 2, a linear regression gives the following: $\Delta G_{Experiments} = 0.17 * \Delta G_{MM-PBSA} - 3.18$. This relationship can then be used with the 34 new complexes to obtain predicted binding free energies which can be compared with experimental ones. The prediction error is very low: RMSD=1.7 kcal/mol. Moreover, the correlation coefficient calculated only on the 34 complexes is R=0.79 (R^2=0.62). Thus, this protocol is transferrable to new BACE-1 ligands.

Conclusions

In the present paper, we have used more than 35 μs of molecular dynamics simulations with MM-PBSA calculations to compute the affinity of ligands to the BACE-1 enzyme receptor which is an important putative target against Alzheimer's disease. First, we used this protocol to study the protonation state of the protein (and not a protocol using structural or docking information), and show that a mono-protonated dyad (on ASP32) is the most likely due to highest Pearson correlation with experiments and lower errors. We then investigated the optimal set of parameters for the MD simulations that leads to the highest correlation coefficient with experiments on an initial set of 12 complexes with diverse types of ligands, and found that correlation coefficients as high as R^2=0.88 can be obtained. We then extended the study to 34 new complexes and found that the parameters which were chosen to maximize the correlation for the initial set of 12 complexes are still valid on this extended set since an overall correlation coefficient of R^2=0.71 is obtained. We also showed that when the linear relation between calculated and experimental binding free energies for 12 complexes is used as a predicting tool, errors as low as

1.7 kcal/mol are obtained for 34 new complexes. Thus, the optimal protocol can be used on newly discovered BACE-1 ligands.

We have shown that very good correlations can be obtained with only 10 simulations of 2ns, and performing many short simulations is preferable to one long (even if it is 50ns long). We compared several Amber force fields and found that all of them provide good correlations, the highest one being obtained with Amber03. TIP3P and SPC/E for the water model give similar results and better than other water models deemed to describe the dynamics of the solvent more accurately. Moreover, a minimal distance of 1.0nm between the protein and box edges is large enough to avoid artifacts from periodic boundary conditions, and better correlation is obtained when the ionic strength is set to 0.2M during the simulations (whereas the ionic strength used for MM-PBSA has no influence). This last result illustrates that adding only the minimal amount of ions to neutralize the simulation box is not a good way to model the experimental setup. Finally, the use of the single-trajectory approximation gives best results. It is likely that most of these conclusions are transferrable to other proteins, that is, many short simulations are better than a single long, all modern protein force fields provide similar sampling hence comparable MM-PBSA values and close correlations, TIP3P and SPC/E water models are equivalent, the standard box size is equivalent to larger ones and a non-zero ionic strength is preferable. Indeed, similar conclusions have already been obtained previously. Crystallographic structures were used as starting point of the simulations to optimize the protocol, however, since docking programs can now provide reliable structures this protocol is transferrable to new ligands discovered during drug design efforts.

Whether or not MM-PBSA will be working for a given class of protein is a question more difficult to answer. MM-PBSA results are system-dependent, and high correlations with

experiments cannot be obtained for certain class of proteins. Knowing it in advance is difficult, but possible sources of failures are the presence of water molecules in the binding site[86] or the change in flexibility with different ligands[87]. In the present case of BACE-1 we obtained a very high correlation, and it must be pointed out that it is especially impressive that there is a correlation of R=0.84 over 8 orders of magnitude of K_d (from 0.017nm to 2000μM i.e. with $\Delta G_{Binding}$ from -14.7 to -3.7 kcal/mol). Many approaches (including the simple ones such as scoring functions) are able to discriminate between very good and very poor binders; however, the method used here can also discriminate ligands within the group of good binders and within the group of poor binders. Thus, the correlation doesn't have a characteristic sigmoid shape of low-resolution methods where all good binders (respectively poor binders) would have a similar predicted binding free energy. Having a linear correlation is important since it allows a smoother implementation in drug design protocols.

The R=0.84 correlation is similar (even slightly higher) to the one obtained with the more resource-demanding FEP on a given family of ligands[10]. Higher correlations have been obtained with MM-PBSA (see for example [88,89]), but always on a small set of complexes and never with more than 45 ligands (to the best of our knowledge). Moreover, the ligand set that was used here is very diverse, going from 22 to 136 atoms, with a molecular mass from 138 to 937 g/mol, representing both small organic fragments and peptide inhibitors (1fkn and 1m4h). Thus, this methodology can be used to select candidates that are good binders found after virtual screening or *de novo* design.

Supporting Information

Description of the MM-PBSA approach. Discussion on the averaging. Distributions of the MM-PBSA energies. Scatter plots corresponding to Table2. Structure of the ligands. Raw MM-PBSA data with uncertainties.

Acknowledgments

The computations in this paper were run on the Odyssey cluster supported by the FAS Research Computing Group (Division of Science, Harvard University).

References

[1] J. Hardy, D. J. Selkoe, *Science* **2002**, *297*, 353–356.
[2] LaFerla, "Amyloid-β and tau in Alzheimer's disease," can be found under http://www.nature.com/nrn/posters/ad/index.html, **2008**.
[3] R. Vassar, *Science* **1999**, *286*, 735–741.
[4] S. L. Cole, R. Vassar, *Mol. Neurodegener.* **2007**, *2*, 22.
[5] G. Probst, Y. Xu, *Expert Opin. Ther. Pat.* **2012**, *22*, 511–540.
[6] R. Yan, R. Vassar, *Lancet Neurol.* **2014**, *13*, 319–329.
[7] M. E. Kennedy, A. W. Stamford, X. Chen, K. Cox, J. N. Cumming, M. F. Dockendorf, M. Egan, L. Ereshefsky, R. A. Hodgson, L. A. Hyde, et al., *Sci. Transl. Med.* **2016**, *8*, 363ra150-363ra150.
[8] Y. Li, L. Han, Z. Liu, R. Wang, *J. Chem. Inf. Model.* **2014**, *54*, 1717–1736.
[9] M. K. Gilson, H.-X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
[10] M. Ciordia, L. Pérez-Benito, F. Delgado, A. A. Trabanco, G. Tresadern, *J. Chem. Inf. Model.* **2016**, *56*, 1856–1871.
[11] N. Homeyer, H. Gohlke, *Mol. Inform.* **2012**, *31*, 114–122.
[12] S. Genheden, U. Ryde, *Expert Opin. Drug Discov.* **2015**, *10*, 449–461.
[13] T. Hou, J. Wang, Y. Li, W. Wang, *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
[14] L. Xu, H. Sun, Y. Li, J. Wang, T. Hou, *J. Phys. Chem. B* **2013**, *117*, 8408–8421.
[15] H. Sun, Y. Li, S. Tian, L. Xu, T. Hou, *Phys. Chem. Chem. Phys.* **2014**, *16*, 16719–16729.
[16] D. P. Oehme, R. T. C. Brownlee, D. J. D. Wilson, *J. Comput. Chem.* **2012**, *33*, 2566–2580.
[17] P.-C. Su, C.-C. Tsai, S. Mehboob, K. E. Hevener, M. E. Johnson, *J. Comput. Chem.* **2015**, *36*, 1859–1873.
[18] C. Wang, P. H. Nguyen, K. Pham, D. Huynh, T.-B. N. Le, H. Wang, P. Ren, R. Luo, *J. Comput. Chem.* **2016**, *37*, 2436–2446.
[19] A. Weis, K. Katebzadeh, P. Söderhjelm, I. Nilsson, U. Ryde, *J. Med. Chem.* **2006**, *49*, 6596–6606.
[20] Y. Li, Z. Liu, J. Li, L. Han, J. Liu, Z. Zhao, R. Wang, *J. Chem. Inf. Model.* **2014**, *54*, 1700–1716.
[21] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[22] A. Šali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.

[23] R. L. Dunbrack, *Curr. Opin. Struct. Biol.* **2002**, *12*, 431–440.
[24] L. Hong, G. Koelsch, X. Lin, S. Wu, S. Terzyan, A. K. Ghosh, X. C. Zhang, T. Jordan, *Science* **2000**, *290*, 150–153.
[25] R. T. Turner, G. Koelsch, L. Hong, P. Castenheira, A. Ghosh, J. Tang, *Biochemistry (Mosc.)* **2001**, *40*, 10001–10006.
[26] L. Hong, R. T. Turner, G. Koelsch, D. Shin, A. K. Ghosh, J. Tang, *Biochemistry (Mosc.)* **2002**, *41*, 10963–10967.
[27] A. K. Ghosh, N. Kumaragurubaran, L. Hong, H. Lei, K. A. Hussain, C.-F. Liu, T. Devasamudram, V. Weerasena, R. Turner, G. Koelsch, et al., *J. Am. Chem. Soc.* **2006**, *128*, 5310–5311.
[28] A. K. Ghosh, N. Kumaragurubaran, L. Hong, S. S. Kulkarni, X. Xu, W. Chang, V. Weerasena, R. Turner, G. Koelsch, G. Bilcer, et al., *J. Med. Chem.* **2007**, *50*, 2399–2407.
[29] E. W. Baxter, K. A. Conway, L. Kennis, F. Bischoff, M. H. Mercken, H. L. De Winter, C. H. Reynolds, B. A. Tounge, C. Luo, M. K. Scott, et al., *J. Med. Chem.* **2007**, *50*, 4261–4264.
[30] U. Iserloh, J. Pan, A. W. Stamford, M. E. Kennedy, Q. Zhang, L. Zhang, E. M. Parker, N. A. McHugh, L. Favreau, C. Strickland, et al., *Bioorg. Med. Chem. Lett.* **2008**, *18*, 418–422.
[31] Y.-S. Wang, C. Strickland, J. H. Voigt, M. E. Kennedy, B. M. Beyer, M. M. Senior, E. M. Smith, T. L. Nechuta, V. S. Madison, M. Czarniecki, et al., *J. Med. Chem.* **2010**, *53*, 942–950.
[32] Z. Zhu, Z.-Y. Sun, Y. Ye, J. Voigt, C. Strickland, E. M. Smith, J. Cumming, L. Wang, J. Wong, Y.-S. Wang, et al., *J. Med. Chem.* **2010**, *53*, 951–965.
[33] J. N. Cumming, E. M. Smith, L. Wang, J. Misiaszek, J. Durkin, J. Pan, U. Iserloh, Y. Wu, Z. Zhu, C. Strickland, et al., *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444–2449.
[34] A. W. Stamford, J. D. Scott, S. W. Li, S. Babu, D. Tadesse, R. Hunter, Y. Wu, J. Misiaszek, J. N. Cumming, E. J. Gilbert, et al., *ACS Med. Chem. Lett.* **2012**, *3*, 897–902.
[35] A. K. Ghosh, K. Venkateswara Rao, N. D. Yadav, D. D. Anderson, N. Gavande, X. Huang, S. Terzyan, J. Tang, *J. Med. Chem.* **2012**, *55*, 9195–9207.
[36] W. Yang, W. Lu, Y. Lu, M. Zhong, J. Sun, A. E. Thomas, J. M. Wilkinson, R. V. Fucini, M. Lam, M. Randal, et al., *J. Med. Chem.* **2006**, *49*, 839–842.
[37] A. Kuglstatter, M. Stahl, J.-U. Peters, W. Huber, M. Stihle, D. Schlatter, J. Benz, A. Ruf, D. Roth, T. Enderle, et al., *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1304–1307.
[38] H. Park, K. Min, H.-S. Kwak, K. D. Koo, D. Lim, S.-W. Seo, J.-U. Choi, B. Platt, D.-Y. Choi, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2900–2904.
[39] C. Björklund, H. Adolfsson, K. Jansson, J. Lindberg, L. Vrang, A. Hallberg, Å. Rosenquist, B. Samuelsson, *Bioorg. Med. Chem.* **2010**, *18*, 1711–1723.
[40] J. Cumming, S. Babu, Y. Huang, C. Carrol, X. Chen, L. Favreau, W. Greenlee, T. Guo, M. Kennedy, R. Kuvelkar, et al., *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2837–2842.
[41] Y. Cheng, T. C. Judd, M. D. Bartberger, J. Brown, K. Chen, R. T. Fremeau, D. Hickman, S. A. Hitchcock, B. Jordan, V. Li, et al., *J. Med. Chem.* **2011**, *54*, 5836–5857.
[42] I. V. Efremov, F. F. Vajdos, K. A. Borzilleri, S. Capetta, H. Chen, P. H. Dorff, J. K. Dutra, S. W. Goldstein, M. Mansour, A. McColl, et al., *J. Med. Chem.* **2012**, *55*, 9069–9088.
[43] S. Yonezawa, K. Fujiwara, T. Yamamoto, K. Hattori, H. Yamakawa, C. Muto, M. Hosono, Y. Tanaka, T. Nakano, H. Takemoto, et al., *Bioorg. Med. Chem.* **2013**, *21*, 6506–6522.
[44] F. Jeppsson, S. Eketjall, J. Janson, S. Karlstrom, S. Gustavsson, L.-L. Olsson, A.-C. Radesater, B. Ploeger, G. Cebers, K. Kolmodin, et al., *J. Biol. Chem.* **2012**, *287*, 41245–41257.

[45] S. W. Gerritz, W. Zhai, S. Shi, S. Zhu, J. H. Toyn, J. E. Meredith, L. G. Iben, C. R. Burton, C. F. Albright, A. C. Good, et al., *J. Med. Chem.* **2012**, *55*, 9208–9223.
[46] M. Mandal, Z. Zhu, J. N. Cumming, X. Liu, C. Strickland, R. D. Mazzola, J. P. Caldwell, P. Leach, M. Grzelak, L. Hyde, et al., *J. Med. Chem.* **2012**, *55*, 9331–9345.
[47] J. P. Caldwell, R. D. Mazzola, J. Durkin, J. Chen, X. Chen, L. Favreau, M. Kennedy, R. Kuvelkar, J. Lee, N. McHugh, et al., *Bioorg. Med. Chem. Lett.* **2014**, *24*, 5455–5459.
[48] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
[49] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, et al., *Bioinformatics* **2013**, *29*, 845–854.
[50] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19–25.
[51] S. Páll, M. J. Abraham, C. Kutzner, B. Hess, E. Lindahl, in *Solving Softw. Chall. Exascale* (Eds.: S. Markidis, E. Laure), Springer International Publishing, Cham, **2015**, pp. 3–27.
[52] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., *Gaussian 09*, Gaussian, Inc., Wallingford, CT, USA, **2009**.
[53] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.
[54] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes* **2012**, *5*, 367.
[55] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*, 247–260.
[56] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, et al., *J. Comput. Chem.* **2003**, *24*, 1999–2012.
[57] G. Bussi Zykova-Timan, T. ..Parrinello,M., *J. Chem. Phys.* **2009**, *130*, 74101.
[58] H. J. C. Berendsen Postma, J. P. M. ..van Gunsteren, W. F. ..DiNola, A.,Haak. J. R., *J. Chem. Phys.* **1984**, *81*, 3684.
[59] M. Parrinello Rahman, A., *J. Appl. Phys.* **1981**, *52*, 7182.
[60] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
[61] B. Hess, *J. Chem. Theory Comput.* **2008**, *4*, 116–122.
[62] W. L. Jorgensen, *J. Am. Chem. Soc.* **1981**, *103*, 335–340.
[63] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[64] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Bioinforma.* **2006**, *65*, 712–725.
[65] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
[66] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **1987**, *91*, 6269–6271.
[67] J. L. F. Abascal, C. Vega, *J. Chem. Phys.* **2005**, *123*, 234505.
[68] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910.
[69] T. Darden York, D. ..Pedersen, L., *J. Chem. Phys.* **1993**, *98*, 10098–10092.
[70] D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, et al., *AMBER 12*, University Of California, San Francisco, **2012**.
[71] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.

[72] T. Hou, R. Yu, *J. Med. Chem.* **2007**, *50*, 1177–1188.
[73] G. Rastelli, A. D. Rio, G. Degliesposti, M. Sgobba, *J. Comput. Chem.* **2010**, *31*, 797–810.
[74] N. Chéron, N. Jasty, E. I. Shakhnovich, *J. Med. Chem.* **2016**, *59*, 4171−4188.
[75] H. Park, S. Lee, *J. Am. Chem. Soc.* **2003**, *125*, 16416–16422.
[76] T. Polgár, G. M. Keserü, *J. Med. Chem.* **2005**, *48*, 3749–3755.
[77] D. A. Sabbah, H. A. Zhong, *J. Mol. Graph. Model.* **2016**, *68*, 206–215.
[78] R. Rajamani, C. H. Reynolds, *J. Med. Chem.* **2004**, *47*, 5159–5166.
[79] F. Sussman, J. M. Otero, M. C. Villaverde, M. Castro, J. L. Domínguez, L. González-Louro, R. J. Estévez, J. C. Estévez, *J. Med. Chem.* **2011**, *54*, 3081–3085.
[80] T. Polgár, C. Magyar, I. Simon, G. M. Keserü, *J. Chem. Inf. Model.* **2007**, *47*, 2366–2373.
[81] A. Barman, R. Prabhakar, *J. Chem. Inf. Model.* **2012**, *52*, 1275–1287.
[82] D. W. Wright, B. A. Hall, O. A. Kenway, S. Jha, P. V. Coveney, *J. Chem. Theory Comput.* **2014**, *10*, 1228–1241.
[83] S. Genheden, U. Ryde, *J. Comput. Chem.* **2009**, 837–846.
[84] P. Mark, L. Nilsson, *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
[85] C. Vega, J. L. F. Abascal, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663.
[86] T. Hou, J. Wang, Y. Li, W. Wang, *J. Comput. Chem.* **2011**, *32*, 866–877.
[87] P. A. Greenidge, C. Kramer, J.-C. Mozziconacci, R. M. Wolf, *J. Chem. Inf. Model.* **2013**, *53*, 201–209.
[88] B. Kuhn, P. A. Kollman, *J. Med. Chem.* **2000**, *43*, 3786–3791.
[89] S. Huo, J. Wang, P. Cieplak, P. A. Kollman, I. D. Kuntz, *J. Med. Chem.* **2002**, *45*, 1412–1419.

Captions

Figure 1. Structure of the BACE-1 enzyme. A snapshot obtained after clustering one of the simulations for the 3kmx structure has been used.

Figure 2. Correlation between computational and experimental binding free energies for the 46 complexes. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, I=0.2M for simulations, I=0.2M for MM-PBSA.

Table 1. List of BACE-1 complexes used in this study. Length is the number of residues in the protein. Charge is the ligand's charge. Mass is the ligand's mass in g/mol. Atoms is the number of atoms in the ligand. K_d is the constant dissociation in nM. Resolution is in Å. Initial set identifies the complexes used to find the optimal parameters. Ref. is the original reference for the crystallographic structure and the K_d measurements.

Table 2. Correlation (R^2) for different states of the BACE-1 protein, calculated on different blocks of the simulation. Errors are in kcal/mol and correspond to the 1-2ns block. State1=ASP32 and ASP228 are not protonated. State2=only ASP32 is protonated. State3=only ASP228 is protonated. State4=both ASP32 and ASP228 are protonated. Best=the lowest binding energy from the four states is kept. "Snapshots" is the number of snapshots used per complex and per state. Parameters: 50 simulations, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

Table 3. Average correlation and error (calculated as the Student's 95% confidence interval) obtained after 100 random extractions of N values. Parameters: State2 for the protonation, simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

Table 4. Comparison of different protein force fields. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

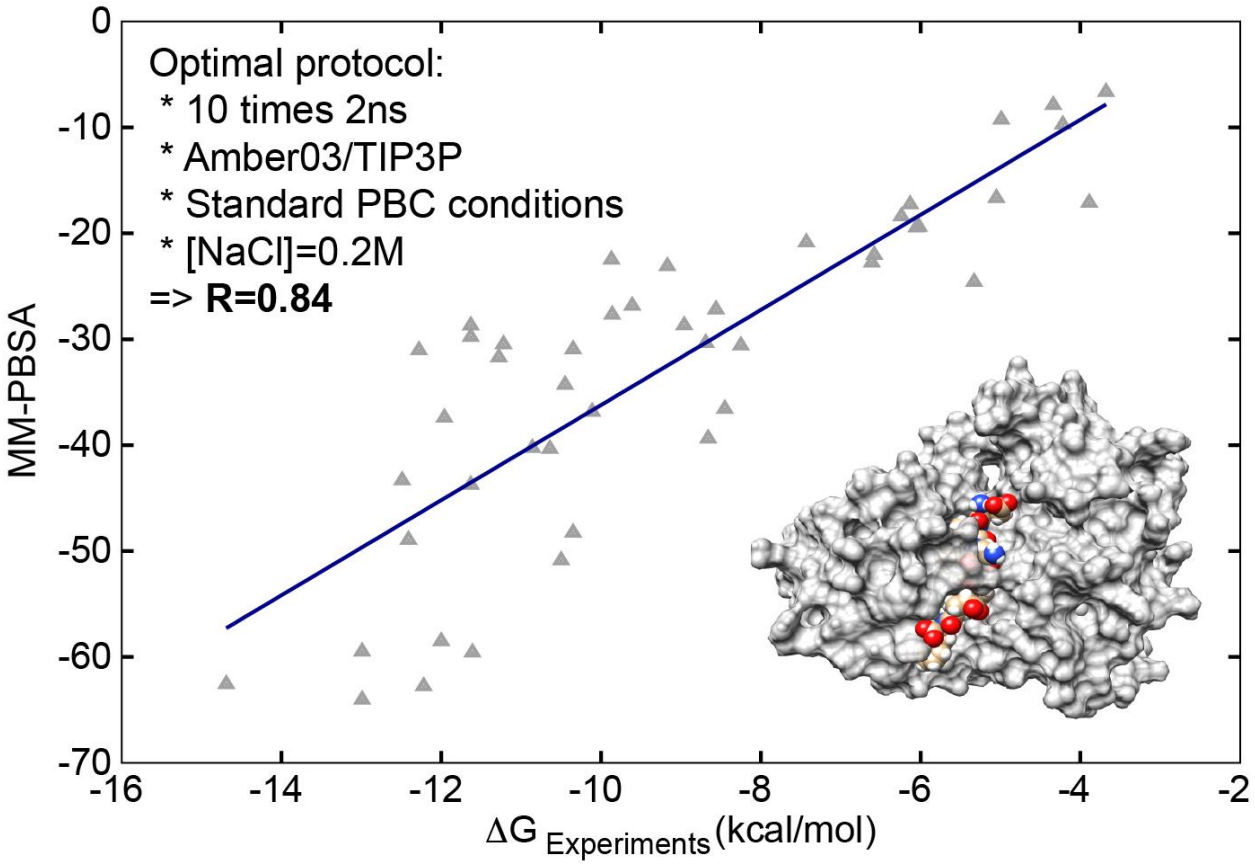
Table 5. Comparison of different water models. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA. Performance is the average performance over all the simulations in ns/day.

Table 6. Comparison of different box sizes and cutoff values. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal number of ions for simulations, I=0 for MM-PBSA. Number of atoms is the average over the 12 complexes.

Table 7. Influence of the ions on the correlation (R^2). Errors (in kcal/mol) are provided for the corresponding parameters for simulations and MM-PBSA (i.e for example, [NaCl]=0.2M for simulations and I=0.2M for MM-PBSA). Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges.

Table 8. Comparison between single and multiple trajectories. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

Graphical Table of Contents



A putative target protein to cure Alzheimer's disease is the BACE-1 enzyme. We searched for a protocol to obtain the most accurate prediction of binding free energy that wouldn't be at the same time too costly in computer resources. Overall, we found a correlation of R=0.84 for 46 ligands, spanning 8 orders of magnitude of K_d with a ligand size from 22 to 136 atoms. This setup is thus generally applicable.

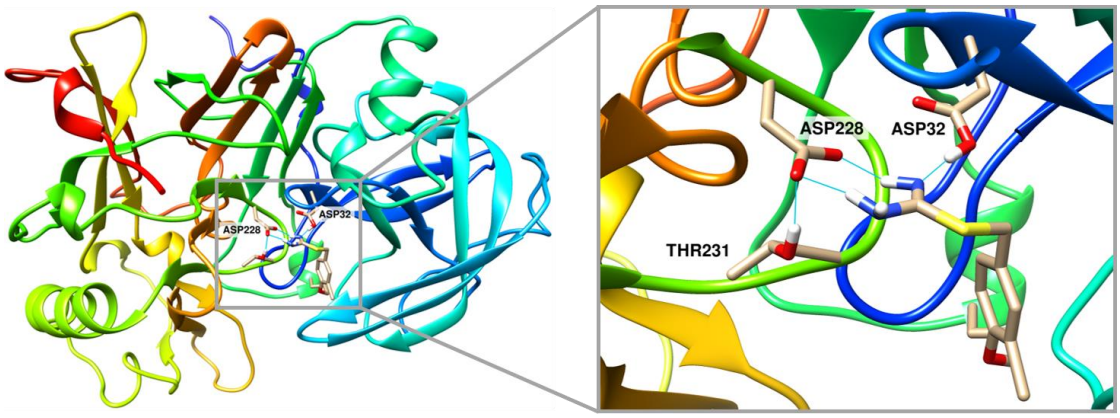


Figure 1. Structure of the BACE-1 enzyme. A snapshot obtained after clustering one of the simulations for the 3kmx structure has been used.

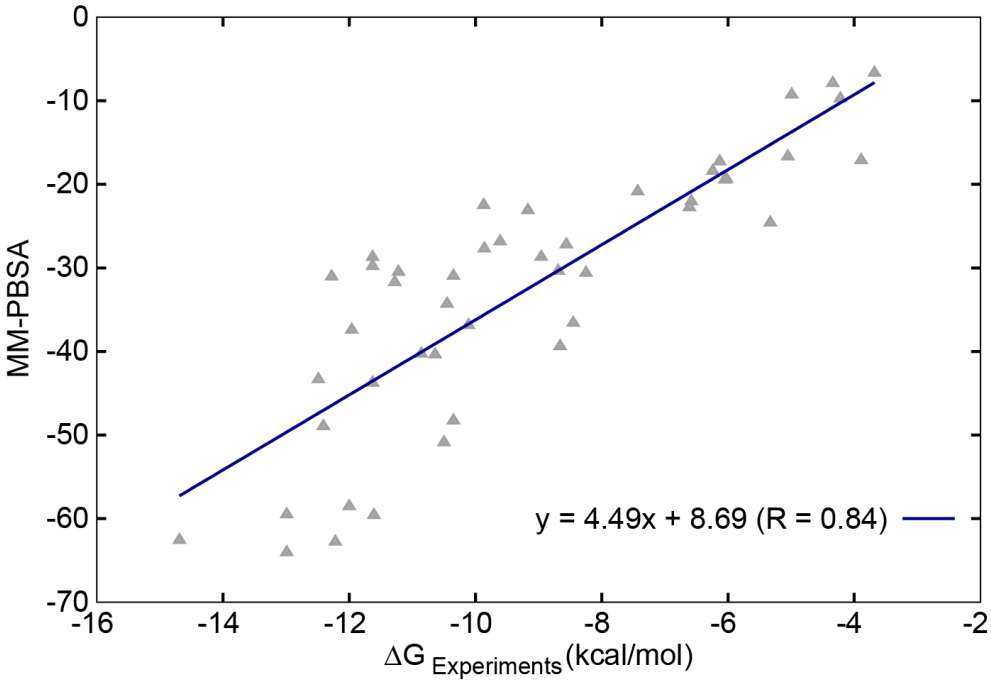
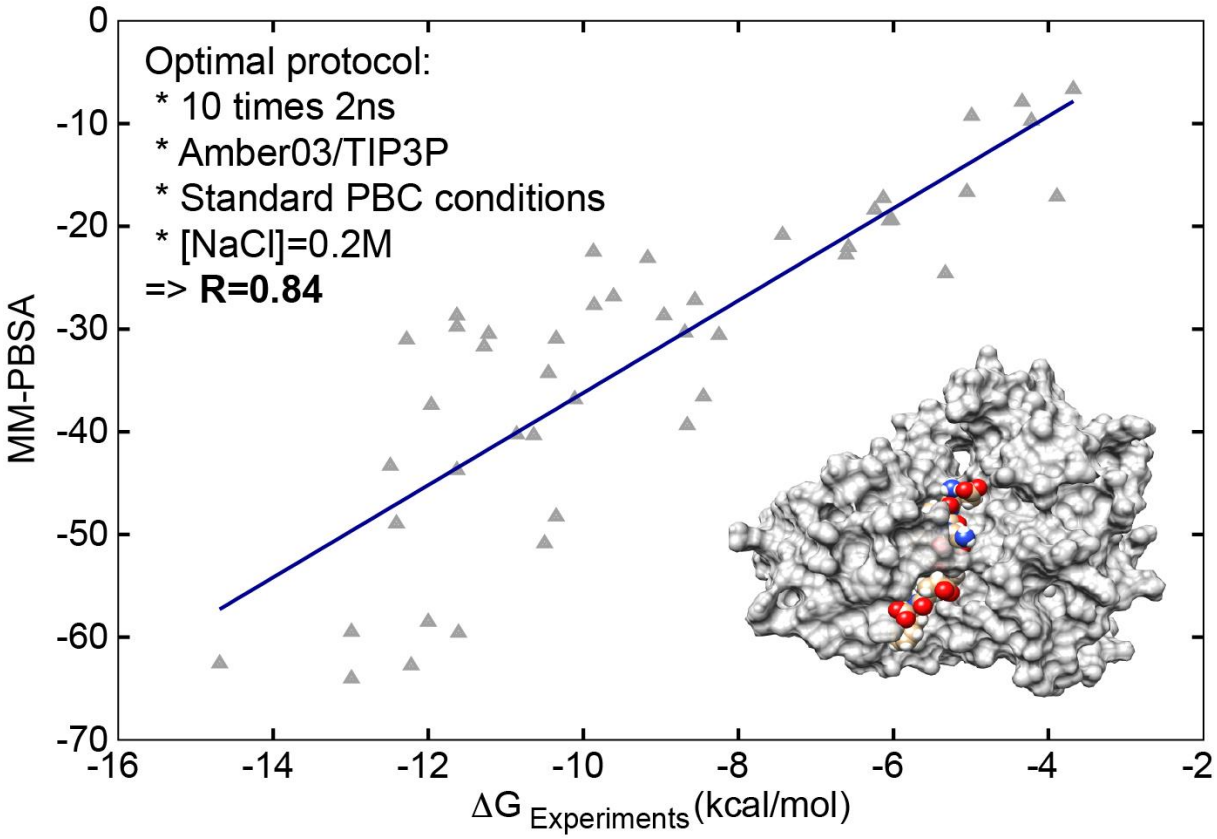


Figure 2. Correlation between computational and experimental binding free energies for the 46 complexes. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, I=0.2M for simulations, I=0.2M for MM-PBSA.

Graphical Table of Contents



A putative target protein to cure Alzheimer's disease is the BACE-1 enzyme. We searched for a protocol to obtain the most accurate prediction of binding free energy that wouldn't be at the same time too costly in computer resources. Overall, we found a correlation of R=0.84 for 46 ligands, spanning 8 orders of magnitude of K_d with a ligand size from 22 to 136 atoms. This setup is thus generally applicable.

PDB ID	Length	Charge	Mass	Atoms	K_d (nM)	$\Delta G_{Binding}$ (kcal/mol)	Resolution	Initial set	Ref.
1fkn	391	-2	894	128	1.60	-12.0	1.9	Yes	[24]
1m4h	391	-3	937	136	0.30	-13.0	2.1	Yes	[25,26]
2g94	389	0	659	99	0.30	-13.0	1.9	Yes	[27]
2p4j	389	0	702	104	1.10	-12.2	2.5	Yes	[28]
2q15	385	0	503	79	11	-10.9	2.4	Yes	[29]
2qmg	390	1	609	84	0.70	-12.5	1.89	Yes	[30]
3kmx	395	0	274	35	15000	-6.6	1.7	Yes	[31]
3kmy	395	0	233	29	32000	-6.1	1.9	Yes	[31]
3l59	414	0	252	31	200000	-5.1	2.0	Yes	[32]
4djx	390	0	341	41	59	-9.9	1.5	Yes	[33]
4fs4	390	0	324	45	270	-9.0	1.7	Yes	[34]
4gid	388	1	667	95	0.017	-14.7	2.0	Yes	[35]
2fdp	388	1	562	83	26	-10.4	2.5	No	[36]
2q11	388	0	421	63	900	-8.3	2.4	No	[29]
3bra	391	1	138	22	2000000	-3.7	2.3	No	[37]
3buf	391	1	152	25	800000	-4.2	2.3	No	[37]
3bug	388	1	166	28	660000	-4.3	2.5	No	[37]
3buh	391	1	220	38	220000	-5.0	2.3	No	[37]
3ckp	388	1	547	80	449	-8.7	2.3	No	[38]
3i25	388	0	896	118	3.1	-11.6	2.1	No	[39]
3l5b	390	0	294	40	123000	-5.3	1.8	No	[32]
3lpi	390	1	658	89	3	-11.6	2.1	No	[40]
3lpk	390	1	686	91	0.8	-12.4	1.9	No	[40]
3rsx	387	0	226	26	38800	-6.0	2.5	No	[41]
3ru1	387	0	311	48	14200	-6.6	2.3	No	[41]
3udh	400	1	189	27	1400000	-3.9	1.7	No	[42]
3wb4	387	0	245	37	36600	-6.1	2.3	No	[43]
3wb5	387	0	259	40	26800	-6.2	2.5	No	[43]
4b05	389	0	431	48	26.1	-10.4	1.8	No	[44]
4dju	390	0	265	35	3600	-7.4	1.8	No	[33]
4djv	390	0	371	49	190	-9.2	1.7	No	[33]
4djw	390	0	342	44	530	-8.6	1.9	No	[33]
4djy	390	0	344	46	5.4	-11.3	1.9	No	[33]
4frs	390	0	375	44	1.7	-12.0	1.7	No	[34]
4fsl	387	0	486	57	20	-10.5	2.5	No	[45]
4h1e	390	0	456	54	3	-11.6	1.9	No	[46]
4h3f	390	0	459	55	1	-12.3	1.7	No	[46]
4h3g	390	0	454	52	6	-11.2	1.9	No	[46]
4h3i	390	0	459	55	3	-11.6	2.0	No	[46]
4h3j	390	0	446	52	90	-9.6	1.6	No	[46]

4ha5	390	0	324	39	58	-9.9	1.8	No	[46]
4r8y	390	0	417	70	640	-8.5	1.9	No	[47]
4r91	390	1	404	72	430	-8.7	1.6	No	[47]
4r92	391	0	440	69	39	-10.1	1.7	No	[47]
4r93	389	0	454	72	16	-10.6	1.7	No	[47]
4r95	390	0	462	73	22	-10.5	2.0	No	[47]

Table 1. List of BACE-1 complexes used in this study. Length is the number of residues in the protein. Charge is the ligand's charge. Mass is the ligand's mass in g/mol. Atoms is the number of atoms in the ligand. K_d is the constant dissociation in nM. Resolution is in Å. Initial set identifies the complexes used to find the optimal parameters. Ref. is the original reference for the crystallographic structure and the K_d measurements.

	State1	State2	State3	State4	BestState	Snapshots
R^2(1-2ns)	0.73	0.86	0.70	0.67	0.85	5000
R^2(2-3ns)	0.72	0.85	0.70	0.68	0.85	5000
R^2(3-4ns)	0.70	0.86	0.70	0.67	0.85	5000
R^2(4-5ns)	0.70	0.85	0.69	0.66	0.83	5000
R^2(1-5ns)	0.71	0.86	0.70	0.67	0.85	20000
RMSD	29.1	35.1	34.9	30.1	37.8	
cRMSD	13.6	15.4	16.7	15.1	15.2	
SD	8.8	7.1	10.9	10.4	7.2	
cSD	2.5	1.8	2.8	3.0	1.7	

Table 2. Correlation (R^2) for different states of the BACE-1 protein, calculated on different blocks of the simulation. Errors are in kcal/mol and correspond to the 1-2ns block. State1=ASP32 and ASP228 are not protonated. State2=only ASP32 is protonated. State3=only ASP228 is protonated. State4=both ASP32 and ASP228 are protonated. Best=the lowest binding energy from the four states is kept. “Snapshots” is the number of snapshots used per complex and per state. Parameters: 50 simulations, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

	N=10	N=20	N=30	N=40	N=50	N=60	N=70	N=80	N=90	N=100
$<R^2>$	0.85	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
Error	0.004	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001

Table 3. Average correlation and error (calculated as the Student's 95% confidence interval) obtained after 100 random extractions of N values. Parameters: State2 for the protonation, simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

	Amber03	Amber99SB	Amber14SB
R^2	0.86	0.84	0.81
RMSD	35.1	35.8	36.8
cRMSD	15.4	16.0	17.2
SD	7.1	7.7	9.0
cSD	1.8	1.9	2.2

Table 4. Comparison of different protein force fields. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.

	TIP3P	SPC/E	TIP4P	TIP4P-2005	TIP5P
R^2	0.86	0.85	0.79	0.80	0.84
RMSD	35.1	34.7	33.5	32.5	34.8
cRMSD	15.4	15.2	14.5	14.0	16.0
SD	7.1	7.1	8.2	7.8	7.8
cSD	1.8	1.8	2.1	2.1	2.0
Performance	29.6	29.3	21.6	21.7	17.1

Table 5. Comparison of different water models. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA. Performance is the average performance over all the simulations in ns/day.

	Small	Medium	Large
Minimal distance between protein and box edges	1.0 nm	1.2 nm	1.4 nm
Force field cutoff	8 Å	10 Å	12 Å
R^2	0.86	0.86	0.87
RMSD	35.1	34.9	35.6
cRMSD	15.4	15.7	15.8
SD	7.1	7.2	7.0
cSD	1.8	1.8	1.7
Number of atoms	54520	61700	70250

Table 6. Comparison of different box sizes and cutoff values. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal number of ions for simulations, I=0 for MM-PBSA. Number of atoms is the average over the 12 complexes.

	[NaCl]=0.0M (MD)	[NaCl]=0.1M (MD)	[NaCl]=0.2M (MD)	[NaCl]=0.3M (MD)
I=0.0M (MM-PBSA)	0.86	0.86	0.87	0.85
I=0.1M (MM-PBSA)	0.86	0.86	0.88	0.85
I=0.2M (MM-PBSA)	0.86	0.86	0.88	0.85
I=0.3M (MM-PBSA)	0.86	0.86	0.88	0.85
RMSD	35.1	35.4	35.0	35.0
cRMSD	15.4	16.0	15.6	15.9
SD	7.1	7.2	6.8	7.4
cSD	1.8	1.8	1.7	1.8

Table 7. Influence of the ions on the correlation (R^2). Errors (in kcal/mol) are provided for the corresponding parameters for simulations and MM-PBSA (i.e for example, [NaCl]=0.2M for simulations and I=0.2M for MM-PBSA). Parameters: State2 for the protonation, 50 simulations of 2ns, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges.

	Single trajectory	Multiple (3) trajectories	Multiple (2) trajectories, ligand from the complex trajectory	Multiple (2) trajectories, protein from the complex trajectory
R^2	0.86	0.11	0.10	0.83
RMSD	35.5	82.8	80.6	43.5
cRMSD	15.7	60.7	65.0	15.9
SD	7.3	60.9	65.5	7.9
cSD	1.8	9.5	11.8	1.6

Table 8. Comparison between single and multiple trajectories. Errors are in kcal/mol. Parameters: State2 for the protonation, 50 simulations, Amber03, TIP3P, minimal distance of 1.0 nm between protein and box edges, minimal number of ions for simulations, I=0 for MM-PBSA.