

MM-GB(PB)SA Calculations of Protein-Ligand Binding Free Energies

Joseph M. Hayes¹ and Georgios Archontis²

*¹Institute of Organic & Pharmaceutical Chemistry
National Hellenic Research Foundation, Athens*

²Department of Physics, University of Cyprus, Nicosia

¹Greece

²Cyprus

1. Introduction

The importance of computational chemistry in modern scientific research is well established. Continuous improvement in software and algorithms for the modeling of chemical interactions has transformed molecular modeling into a powerful tool for many current day research projects. From a medical perspective, one of the ultimate goals in computer-aided drug design (CADD) is the accurate prediction of ligand-binding affinities to a macromolecular target, which can facilitate and speed the routine identification of new candidates in early stage drug discovery projects (Gilson & Zhou, 2007; Hayes & Leonidas, 2010). In particular, structure-based modeling provides an efficient pathway towards exploiting known three-dimensional structural data in the design and proposal of new molecules for experimental evaluation. Docking calculations are now widely used in high-throughput virtual screening of structurally diverse molecules from available compound libraries/databases against specific targets. Once initial “hits” or “lead” molecules are identified (normally low μM inhibitors), modification of their chemical features in the “lead optimization” phase can improve their binding affinities and fine-tune other desirable drug-like properties. However, docking calculations currently have limited success beyond the lead identification stage, where more accurate lower-throughput computational methods are needed. In this regard, the Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) and Molecular Mechanics/Poisson-Boltzmann Surface Area (MM-PBSA) methods calculate binding free energies using molecular mechanics (forcefields) and continuum (implicit) solvation models (Kollman et al., 2000). They have been successfully applied across a range of targets and are implemented in software programs such as Amber (Case et al., 2005), Delphi (Rocchia et al., 2001) and Schrödinger (Du et al., 2011). With a target readership from beginner to expert, the current chapter provides an extensive and critical overview of MM-GB(PB)SA methods and their applications. The theoretical foundation of the MM-GB(PB)SA method is first described. We then discuss key aspects which improve the accuracy of results, and highlight potential caveats due to the approximations inherent in the methods. The chapter concludes with a review of recent representative applications, which illustrate both successes and limitations. The emphasis of this chapter is on structure-

based drug design (SBDD) efforts, however, the methods have widespread applicability in other areas such as in supramolecular chemistry.

2. Theoretical background – “Pathway” and “Endpoint” methods

Low-throughput computational approaches for the calculation of ligand binding free energies can be divided into “pathway” and “endpoint” methods (Deng & Roux, 2009; Gilson & Zhou, 2007). In pathway methods, the system is converted from one state (e.g., the complex) to the other (e.g., the unbound protein/ligand). This can be achieved by introducing a set of finite or infinitesimal “alchemical” changes to the energy function (the Hamiltonian) of the system through free-energy perturbation (FEP) or thermodynamic integration (TI), respectively (Kollman, 1993; Straatsma & McCammon, 1992). The fundamentals of FEP and TI methods were introduced many decades ago by John Kirkwood (Kirkwood, 1935) and Robert Zwanzig (Zwanzig, 1954). In recent years, their use in the computation of absolute binding affinities has become feasible due to increases in computational power, the development of more accurate models of atomic interactions (Cornell et al., 1995; MacKerell et al., 1998; van Gunsteren et al., 1996), the clarification of the underlying theoretical framework and the introduction of methodological advances (Boresch et al., 2003; Bowers et al., 2006; Deng and Roux, 2009; Gilson et al., 1997; Gilson & Zhou, 2007; Lee & Olson, 2006; Mobley et al., 2007). Combined with atomistic molecular dynamics (MD) or Monte Carlo (MC) simulations in explicit water solvent models, they are arguably the most accurate methods for calculating absolute or relative ligand binding affinities.

The “alchemical” computation of *differences* in binding affinities (rather than *absolute* affinities) among a set of related ligands for the same target protein is more accurate and technically simpler. A thermodynamic cycle illustrating the basic principles is shown in Figure 1 (Tembe & McCammon, 1984). The horizontal legs describe the experimentally accessible actual binding processes, with free energies $\Delta G_{\text{bind}}(L1)$ and $\Delta G_{\text{bind}}(L2)$. Since the free energy is a state function, the relative binding free energy $\Delta\Delta G_{\text{bind}}$ is exactly equal to the difference of the free energies in the horizontal or vertical legs:

$$\Delta\Delta G_{\text{bind}} = \Delta G_{\text{bind}}(L2) - \Delta G_{\text{bind}}(L1) \quad (1)$$

$$= \Delta G_{\text{complex}}(L1 \rightarrow L2) - \Delta G_{\text{free}}(L1 \rightarrow L2) \quad (2)$$

The simulations follow the vertical steps (Eq. (2)) or unphysical processes, by simulations in water solution that gradually change the energy-function of the system from one “endpoint” to the other through a series of intermediate hybrid states. From Figure 1, this involves the stepwise “alchemical” transformation of ligand L1 to L2 both in its ‘free’ state (unbound) and in the bound complex, through gradual changes in the forcefield parameters describing the ligand interactions. This leads to the free energy changes $\Delta G_{\text{free}}(L1 \rightarrow L2)$ and $\Delta G_{\text{complex}}(L1 \rightarrow L2)$, respectively. Averaging over both transformation directions is often used to improve the free-energy estimates, although this is not always the case (Lu & Woolf, 2007). These calculations can be accurate, if conducted with the appropriate care. An overview of current state-of-the-art methods for absolute and relative affinity calculations is in (Chodera et al., 2011).

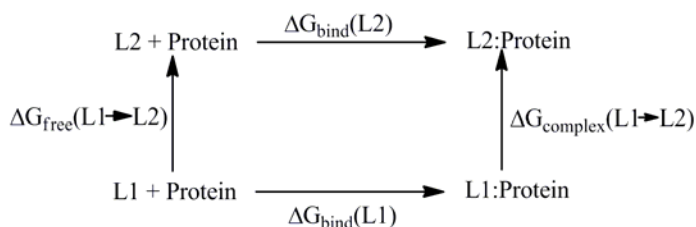


Fig. 1. Thermodynamic cycle linking the binding of two ligands L1 and L2 to a protein in solution.

To conclude, the emerging implementation of biomolecular codes on GPU architectures (Harvey et al., 2009; Stone et al., 2011) and the development of simple free-energy protocols (Boresch & Bruckner, 2011) make atomistic methods of absolute or relative affinities very promising for larger-scale calculations in the near future. Nevertheless, at present they are still relatively time-consuming, and require considerable expertise and planning. They preclude the consideration of more than a few complexes per day on a dedicated CPU cluster with a few tens of nodes. A trade-off between computational expense and accuracy is therefore required when the goal is to investigate and compare the binding strengths of a structurally diverse and/or larger set of ligands via MD simulations. For this purpose, much less computationally demanding „endpoint“ methods are often successfully applied, such as the „linear interaction energy“ (LIE) (Åqvist et al., 2002) or the molecular mechanics – Poisson Boltzmann (MM-PBSA) (Kollman et al., 2000) and the related molecular mechanics – generalised Born (MM-GBSA) approximation (Gohlke et al., 2003). All these methods compute binding free energies along the horizontal legs of Figure 1, but use only models for the “endpoints” (bound and unbound states). The MM-PB(GB)SA methodology is now described in more detail.

3. The MM-GB(PB)SA methodology

Using MM-GBSA and MM-PBSA methods, relative binding affinities for a set of ligands to a given target can often be reproduced with good accuracy and considerable less computational effort compared to full-scale molecular dynamics FEP/TI simulations. Furthermore, free-energies can be decomposed into insightful interaction and desolvation components (Archontis et al., 2001; Hayes et al., 2011; Polydoridis et al., 2007).

3.1 Thermodynamics & calculation framework

In the MM-GB(PB)SA formulation, the binding free energy of a ligand (L) to a protein (P) to form the complex (PL) is obtained as the difference (Pearlman, 2005):

$$\Delta G_{\text{bind}} = G(PL) - G(P) - G(L) \quad (3)$$

The free energy of each of the three molecular systems P, L, and PL is given by the expression:

$$G(X) = E_{\text{MM}}(X) + G_{\text{solv}}(X) - TS(X) \quad (4)$$

In Eq. (4), E_{MM} is the total molecular mechanics energy of molecular system X in the gas phase, G_{solv} is a correction term (solvation free energy) accounting for the fact that X is surrounded by solvent, and S is the entropy of X .

To apply the MM-GB(PB)SA formulation, a representative set of equilibrium conformations for the complex, free protein and free ligand are first obtained by atomistic MD simulations in explicit solvent. In this post-processing phase, the solvent is discarded and replaced by a dielectric continuum. Changes (Δ) in the individual terms (ΔE_{MM} , ΔG_{solv} , $-T\Delta S$) of Eq.(4) between the unbound states and the bound (complex) state are calculated, and contribute to the binding free energies according to Eq.(3). Computation of each of the terms in Eq. (4) is now described in more detail.

E_{MM} is the sum of the bonded (internal), and non-bonded electrostatic and van der Waals energies

$$E_{MM} = E_{bonded} + E_{elec} + E_{vw} \quad (5)$$

These energy contributions are computed from the atomic coordinates of the protein, ligand and complex using the (gas phase) molecular mechanics energy function (or forcefield). The solvation free energy term G_{solv} contains both polar and non-polar contributions. The polar contributions are accounted for by the generalized Born, Poisson, or Poisson-Boltzmann model, and the non-polar are assumed proportional to the solvent-accessible surface area (SASA), *c.f.* Section 3.2:

$$G_{solv} = G_{PB(GB)} + G_{SASA} \quad (6)$$

Finally, the entropy S is decomposed into translational, rotational and vibrational contributions. The first two are computed by standard statistical-mechanical expressions, and the last is typically estimated from a normal-mode (harmonic or quasiharmonic) analysis (Brooks et al., 1995; Karplus & Kushick 1981; Tidor & Karplus 1994). In practice, current software implementations normally determine all three contributions to S as part of a normal-mode analysis.

To improve the accuracy of the computed binding free energies, the various terms of Eq. (4) are averaged over multiple conformations or MD snapshots (typically a few hundred for the E_{MM} and G_{solv} contributions). Depending on the extent of conformational fluctuations in the system under consideration, the convergence into stable values may require relatively long (multi-ns) simulations. The computation of the entropy term, however, requires the extensive minimization of the trajectory conformations for the protein, ligand and complex to local minima on the potential energy surfaces, followed then by normal mode analysis. This procedure is costly and prevents the consideration of a large number of conformations; insufficient sampling can therefore sometimes be an issue. To decrease the computational cost, the protein can be truncated beyond a certain cutoff distance and the system minimized using a distance-dependent dielectric, which simulates the deleted surroundings (Kongsted & Ryde, 2009). However, a large variation of the entropy term often results from these 'free' minimizations. Including a fixed buffer region (with water molecules) beyond the cut-off can lead to more stable entropy predictions (Kongsted & Ryde, 2009).

The internal energy terms (E_{bonded}) of the protein and complex can be on the order of a few thousand kcal/mol, and can introduce large uncertainties in the computed binding free

energies. This is prevented in the “single-trajectory” approximation (Gohlke & Case, 2004; Page & Bates, 2006), which employs simulations of a single state (the complex) to generate conformations for all three states (complex, protein and ligand). For each MD conformation sampled, the resulting internal energy terms of the protein and ligand are identical in the bound and the unbound states and cancel exactly in Eq. (3). Hence, effectively only the protein-ligand (non-bonded) interaction energies of the E_{MM} term in Eq. (5) contribute to ΔG_{bind} . ‘Single-trajectory’ simulations significantly reduce computational effort and are generally sufficiently accurate for most applications. The downside of the approximation is that any explicit structural relaxation of the protein and ligand upon binding is ignored. Although charge reorganization can be partly taken into account implicitly by setting the protein/ligand (internal) dielectric constants (Section 3.2) to values larger than $\epsilon_{in} = 1-2$ (Archontis et al., 2001; Archontis & Simonson, 2005; Schutz & Warshel, 2001; Simonson, 2003), the neglect of explicit structural relaxation may introduce errors depending on the system (Tamamis et al., 2010). Separate MD simulations for the complex, and unbound receptor and ligands may also be performed (the “three-trajectory” approximation) but require greater computational effort, although in theory should yield more accurate results. Indeed, Yang and co-workers have recently shown that including separate simulations for the ligand and accounting for the “ligand reorganization” free energy led to significant improvements in binding affinity predictions for a set of ligands targeting XIAP (Yang et al., 2009). In certain cases, therefore, the added expense of separate simulations may be justified.

3.2 Solvation models – GBSA and PBSA

Proteins function usually inside aqueous solutions or in membrane environments, which are in the vicinity of an aqueous medium. The surrounding solvent can influence protein stability and function, ligand binding and protein-protein association. Since the solvent modifies in a non-trivial manner the intramolecular and intermolecular interactions, an accurate inclusion of solvent effects in biomolecular modeling and simulation is a challenging task.

Currently, the most accurate treatment in molecular simulations is achieved by atomic-detail models that represent explicitly the biomolecule and its surrounding environment. Several water models are used successfully to describe the aqueous environment in atomistic simulations; examples include SPC (Berendsen et al., 1981), SPC/E (Berendsen et al., 1987), TIP3P and TIP4P (Jorgensen et al., 1983), and TIP5P (Mahoney & Jorgensen, 2000). In practice, the explicit inclusion of water leads to a considerable increase in both the size of the simulation system and the computational cost of the simulation itself. Furthermore, the computation of solvation or binding free energies requires an exhaustive sampling of the solvent degrees of freedom. A much less costly approach is to represent the solvent implicitly in the simulation, through the incorporation of additional “potential of mean force” terms (Roux, 2001; Roux & Simonson, 1999) in the gas-phase energy function (e.g., Eq. (7) below). These terms depend only on the atomic coordinates of the solute, and express the solute free energy for a given configuration, after the solvent degrees of freedom have been “integrated out” (Roux, 2001; Roux & Simonson, 1999). Thus, the simulation system has the same number of degrees of freedom as in the gas phase and there is no need for explicit sampling over solvent degrees of freedom. The MM-PB(GB)SA method considered here, combine atomistic simulations in explicit solvent for the generation of representative

biomolecular conformations with an implicit-solvent estimation of the binding free energies, in a post-processing step.

Conceptually, most implicit solvent models decompose the solvation process into three sequential steps (Cramer & Truhlar, 1999): i) creation of a cavity in solution to accommodate the biomolecule; ii) switching-on dispersion interactions between the biomolecule and surrounding medium, while all atomic charges are set to zero; and iii) switching-on the biomolecular charges. The solvation free energies of steps i) and ii) are normally assumed to be proportional to the SASA of the biomolecule and represent the *non-polar* contributions (G_{SASA}) to G_{solv} in Eq. (6), although the validity of this approximation has been questioned for step ii) (Levy et al., 2003). With a positive coefficient of proportionality, an increase in the SASA is associated with an unfavorable increase in solvation free energy, which is partly accounted for by the tendency of non-polar residues to be solvent-excluded. The equation typically used is of the form:

$$G_{SASA}(X) = \gamma \cdot SASA + \beta \quad (7)$$

with the γ and β parameter values dependant on the method and solvation model (PBSA or GBSA) used (Rastelli et al., 2010).

Meanwhile, step iii) calculates the contribution to solvation free energy due to the charge /electrostatic interactions of the solute with the surrounding solvent, the *polar* contributions ($G_{PB(GB)}$) to G_{solv} in Eq. (6). In continuum-electrostatics models such as PB and GB, the solute is treated as a low-dielectric cavity embedded in a high dielectric medium. The solute charges are in the simplest and most common approximation centered on the individual atoms. The resulting solvation free energy of a molecule X is expressed as (Simonson, 2003):

$$G_{PB(GB)}(X) = \frac{1}{2} \sum_{i,j \in X} q_i q_j g_{ij}^{PB(GB)} \quad (8)$$

where the summation is over all the atomic charges $\{q_i\}$. The quantity $g_{ij}^{PB(GB)}$ is determined using the PB model by numerical solution of the Poisson or Poisson-Boltzmann equation (depending on the existence of salt), or using the GB model by an analytical expression with the functional form (Simonson, 2003; Still et al., 1990):

$$g_{ij}^{GB} = \left(\frac{1}{\epsilon} - 1 \right) \left[r_{ij}^n + B_{ij} \exp \left(- \frac{r_{ij}^n}{A B_{ij}} \right) \right]^{-1/n} \quad (9)$$

The parameters B_{ij} depend on the position (distance from the solute-solvent dielectric boundary) of atoms i and j , and the shape of the entire biomolecule; ϵ is the solvent dielectric constant, and r_{ij} is the distance between i and j . The constants n and A were set to $n=2$ and $A=4$ in the original formulation of Still and coworkers (Still et al., 1990).

In the PB model, the solute dielectric constant (ϵ_{in}) affects the computed functions g_{ij}^{PB} and Eq. (8). Meanwhile, in the GB model, the solute dielectric constant drops out from the final expression in Eq.(9), due to the approximations used to arrive at an analytic formula. An ϵ_{in} value other than 1 can still be used; in this case, the first parenthesis on the right-hand side of Eq. (9) becomes $(1/\epsilon - 1/\epsilon_{in})$ and the GB expression yields the free energy of transferring

the solute from an infinite reference medium with dielectric constant ϵ_{in} into solution. A careful discussion of this point is in (Bashford & Case, 2000).

Application of Eq. (8) to a protein:ligand complex (PL) and the dissociated protein (P) and ligand (L) yields the electrostatic (polar) solvation free energy contribution to Eq. (3):

$$\Delta G_{PB(GB)} = G_{PB(GB)}(PL) - G_{PB(GB)}(P) - G_{PB(GB)}(L) \quad (10)$$

An advantage of these methods is that they facilitate the decomposition of the total solvation free energy into insightful components. Indeed, the summation over atomic charges in Eq. (8) implies that the electrostatic free energy of Eq. (10) can be partitioned into interaction and desolvation components (Archontis et al., 2001; Hendsch & Tidor, 1999):

$$\Delta G_{PB(GB)} = \sum_{i \in P, j \in L} q_i g_{ij}^{PL} q_j + \left[\frac{1}{2} \sum_{i, j \in P} q_i g_{ij}^{PL} q_j - \frac{1}{2} \sum_{i, j \in P} q_i g_{ij}^P q_j \right] + \left[\frac{1}{2} \sum_{i, j \in L} q_i g_{ij}^{PL} q_j - \frac{1}{2} \sum_{i, j \in P} q_i g_{ij}^L q_j \right] \quad (11)$$

The notation g_{ij}^x implies that the interaction between charges i and j , as determined by the function g_{ij} , is in general different in the complex, free protein and free ligand. The first term on the right-hand side of Eq. (11) is the "interaction term" and arises from direct interactions between the protein and ligand charges, which are only present in the complex; the next term [the first brace] is a protein "desolvation" term, arising from the replacement of high-dielectric solvent by the low-dielectric ligand in the protein vicinity, as well as structural relaxation and changes in the charge distribution of the protein. The last term [second brace] corresponds to the desolvation of the ligand.

The "interaction term" can be further decomposed into contributions from specific residues (Archontis et al., 2001; Hendsch & Tidor 1999). For example, the contribution of a protein residue R to this term is given by the following expression

$$\Delta G_{PB(GB)}^R = \sum_{i \in R, j \in L} q_i g_{ij}^{PL} q_j \quad (12)$$

These components provide useful insights on the origin of the binding free energy values. They can help interpret differences in binding affinities for a series of related complexes and guide the design of modified proteins or ligands. For example, in a study of amino acid binding to native and mutant aspartyl-tRNA synthetase, residue decomposition identified the protein residues discriminating between the cognate ligand (aspartic acid) and the analogue asparagine (Archontis et al., 2001). In another study of RNase A recognition by dinucleotidic inhibitors, a similar decomposition attributed the stronger binding of the most potent inhibitor to interactions with two active site lysines (Polydoridis et al., 2007).

Finally, Hou and co-workers very recently evaluated the performance of MM-GBSA and MM-PBSA for predicting binding free energies based on molecular dynamics simulations (Hou et al., 2011a). Their results showed that MM-PBSA performed better in calculating absolute binding free energies compared to MM-GBSA but not necessarily for the relative binding free energies, sufficient for most applications in computational drug design. Interestingly, in a study of the accuracy of continuum solvation models for drug-like molecules, GB methods typically were more stable and gave more accurate results than the widely used PB methods (Kongsted et al., 2009).

3.3 Selection of MD trajectory conformations/snapshots

A recent study for the binding of seven biotin analogues to avidin suggested that to obtain statistically converged MM-GBSA results, several independent simulations each with sampling times of 20-200 ps (averaging the results) is more effective than a single long simulation (Genheden & Ryde, 2010). ‘Single-trajectory’ simulations of the complex are generally sufficiently accurate for most applications, and while MD simulation length does have an obvious impact on the accuracy of predictions, longer MD simulations doesn’t necessarily mean better predictions (Hou et al., 2011a). For the calculations of the ΔE_{MM} and ΔG_{solv} terms, a large ensemble (e.g. several hundred) conformations are typically extracted in small intervals from the single MD trajectory of the complex. Alternatively, averaging over a select few receptor-ligand binding conformations from the MD trajectory via clustering has proved effective (Section 4.1), as well as more time efficient (Hayes et al., 2011). MM-GB(PS)SA calculations on single (minimized) structures has also recently been proposed and validated (Kuhn et al., 2005; Rastelli et al., 2010), but not necessarily for structures generated from MD simulations (Section 4.2). Meanwhile, for the entropy term calculated using normal mode analysis, fewer snapshots (typically less than a 100) are employed, due to the computational cost involved. As already highlighted (Section 3.1), a larger number of snapshots may be required for more stable and accurate predictions. These calculations, however, are computationally expensive and often not feasible with limited computational resources. Consequently, neglect of the entropy term can in some cases lead to sufficient or more accurate predictions for ranking of ligand binding affinities in certain macromolecular systems (Hayes et al., 2011; Hou et al., 2011a; Rastelli et al., 2010).

3.4 Limitations & caveats of MM-GB(PB)SA calculations

MM-GB(PB)SA methods are widely recognised as valuable tools in CADD applications. However, as with any method they have limitations and caveats, which need to be considered. First, while useful for ranking relative ligand binding affinities, these methods lack the required accuracy for absolute binding free energy predictions (Hou et al., 2011a; Singh & Warshel, 2010). The inclusion of entropic contributions brings the MM-GB(PB)SA values somewhat closer to experimental absolute affinities (Gilson & Zhou, 2007). However, such entropic terms are costly and contain large uncertainties. Force-field inconsistencies may also be an issue: PB and GB results depend strongly on adequate atomic charges and van der Waals radii, which are often optimized for MD simulations. The MM-GB(PB)SA results may be influenced by system-dependent properties, such as the features of the binding site, the extent of protein and ligand conformational relaxation upon association, and the protein and ligand charge distribution (Kuhn et al., 2005; Hou et al., 2011a). Continuum electrostatics models ignore the molecular structure of the solvent; in some cases this might affect the results, particularly when key receptor-ligand interactions are bridged by water molecules, *c.f.* Section 4.1 (Hayes et al., 2011). Furthermore, the value of the protein/ligand dielectric constant is empirically chosen, and takes into account not only the protein and ligand structural relaxation, but also other error-introducing factors such as the ones mentioned above (Archontis & Simonson, 2005; Schutz & Warshel, 2001). Hou and co-workers suggested in a recent MM-PBSA study that the use of $\epsilon_{in} = 4$ for a highly charged protein-ligand binding interface, $\epsilon_{in} = 2$ for a moderately charged binding interface and $\epsilon_{in} = 1$ for a hydrophobic binding interface may improve ligand ranking (Hou et al., 2011a). The lack of a consistent optimum dielectric constant for MM-PBSA calculations has been noted

by other workers (see for e.g. (Aleksandrov et al. 2010)), although generally a value $\epsilon_{in} = 4$ often gives satisfactory results (Aleksandrov et al., 2010; Archontis et al., 2001; Thompson et al., 2006). As a final note, MM-GB(PB)SA calculations require some degree of user expertise and planning, from the initial set-up and analysis of the MD simulations through to the binding free energy calculations themselves.

3.5 Variations & extensions of MM-GB(PB)SA

While molecular docking algorithms are computationally efficient methods used to screen a large number of ligands against a given target in reasonable time, generating and post-processing MD ensembles for more than a few receptor-ligand structures in MM-GB(PB)SA calculations is currently impractical. Recently, however, MM-GB(PB)SA methods are receiving plaudits as post-docking methods in virtual screening experiments. Post-processing of single docking poses using MM-GB(PS)SA algorithms can improve correlations between predicted and experimental binding affinities with a number of successes reported, *c.f.* for example (Du et al., 2011; Hou et al., 2011b; Lyne et al.; 2006). This is consistent with the previously mentioned discovery that select single receptor-ligand structures can prove as accurate as sampling over large numbers of MD trajectory snapshots in MM-GB(PB)SA applications (Kuhn et al., 2005; Rastelli et al., 2010). Post-docking MM-GBSA is implemented in Schrödinger software using the program Prime, with options to include receptor and ligand flexibility; the entropy term is neglected by default. Other recent extensions of MM-PBSA exploit quantum mechanics (QM) methods in QM/MM-PBSA calculations (Gräter et al., 2005; Manta et al., 2012; Wang & Wong, 2007). Here, a “hybrid” gas phase energy term ($E_{QM/MM}$) effectively replaces the pure molecular mechanics energy (E_{MM}) term in Eq. (5). Representing the ligand by QM has the advantage, for example, to eliminate a frequently encountered problem of deficient ligand forcefield parameters. However, calculations of this sort are significantly more expensive than MM-GB(PB)SA, and therefore are typically only viable for binding predictions on relatively few ligands.

4. Recent applications of MM-GB(PS)SA

In the present section, we will review recent examples of the use of MM-GB(PS)SA calculations for calculating ligand binding free energies highlighting both successes and limitations.

4.1 Example 1: Phosphorylase kinase ATP-binding site inhibitors

With an aim towards glycogenolysis control in type 2 diabetes, indirubin ($IC_{50} > 50 \mu M$), indirubin-3'-oxime ($IC_{50} = 144 nM$), KT5720 ($K_i = 18.4 nM$) and staurosporine ($K_i = 0.37 nM$) were investigated as phosphorylase kinase (PhK γ trnc) ATP-binding site inhibitors (Hayes et al., 2011).

Due to the lack of experimental structural information for binding of these ligands (Figure 2), MD simulations in explicit solvent using Desmond 2.0 (Bowers et al., 2006) were performed for each receptor-inhibitor complex, with 4 ns production runs following an initial equilibration period. A computationally-efficient multiple timestep RESPA integration algorithm was employed with timesteps of 2, 2 and 6 fs for bonded, “near” and “far” non-bonded interactions, respectively. Energy and trajectory atomic coordinate data

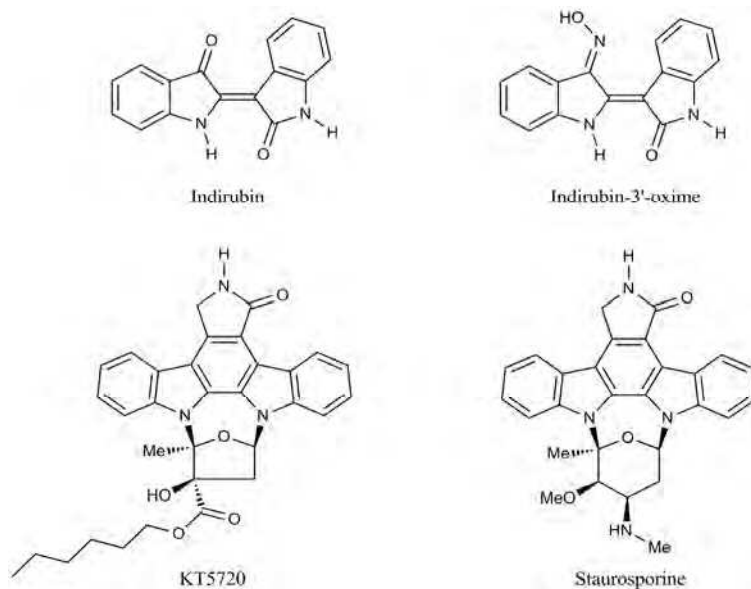


Fig. 2. The ATP-binding site inhibitors of phosphorylase kinase, as studied in Ref. (Hayes et al., 2011).

were recorded every 1.2 and 2.1 ps, respectively. For the trajectory analysis preceding the ‘single-trajectory’ based MM-GBSA calculations, both VMD (Humphrey et al., 1996) and Desmond’s Maestro simulation analysis tools were employed. The extracted MD trajectory binding site conformations (inhibitors + residues within 7 Å) of each complex were clustered into 10 groups based on atomic root-mean-square-distances (RMSDs). Every second frame (snapshot) of the last 3 ns of the production run (analysis phase) was used in the hierarchical clustering algorithm employed by the Desmond Maestro’s Trajectory Clustering module. The representative complex of each of the 10 binding site cluster families (waters and counterions deleted) was then used in MM-GBSA calculations of binding free energies using Eq. (3). Schrödinger software with the MacroModel 9.7 Embrace module was used for the ΔE_{MM} and ΔG_{solv} calculations, while the entropy change, ΔS , was calculated for the minimized representatives using Rigid Rotor Harmonic Oscillator (RRHO) calculations. Using this algorithm, the change in vibrational, rotational and translational (VRT) entropy of the ligands on binding was estimated. Finally, the thermodynamic average ΔG_{bind} values (300 K) were estimated using the corresponding values for the 10 cluster representatives:

$$\Delta G_{bind} = \sum_{i=1}^{10} p_i \Delta G_{bind}(i) \quad (13)$$

The sum i was over the 10 cluster representatives, with p_i defined as the cluster frequency:

$$p_i = \frac{N_i}{N_{total}} \quad (14)$$

where N_i the number of frames in cluster i , and N_{total} the total number of frames.

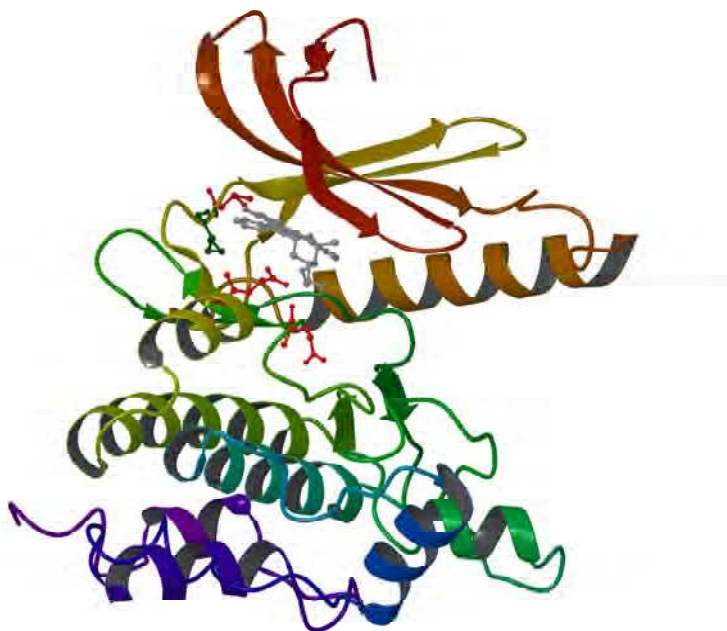


Fig. 3. Predicted binding of staurosporine (gray) at the ATP-binding site of PhKytnc. Key interacting residues are colored by type – polar/charged residues shown in red (D104, E110 and E153) and hydrophobic in green (M106). Hydrogen bonds are formed with M106 and D104 backbones (hinge region), E110 and E153.

The relative binding affinities from the final ΔG_{bind} values were generally in agreement with experiment, except the rankings of KT5720 and staurosporine (representative complex shown in Figure 3) were reversed by 0.7 kcal/mol. The discrepancy was accounted for by neglect of certain key contributions to the binding free energies using the MM-GBSA algorithm. Notably, accounting for and estimating the loss of the conformational entropy for the more flexible KT5720 ligand yielded ΔG_{bind} values 1.2 – 2.6 kcal/mol in favour of staurosporine binding and hence in agreement with the experimental rankings. Further, whereas staurosporine had no key receptor-ligand bridging waters, the MD simulations revealed a key role of bridging waters for KT5720 binding. The entropy loss associated with a bound water molecule in a protein-ligand complex is sometimes important (but typically neglected in MM-GB(PB)SA calculations), with an upper bound free energy cost of 2 kcal/mol at 300 K suggested (Dunitz, 1994).

4.2 Example 2: Interpretation of species specificity of compstatin, a peptidic inhibitor of the complement system

The complement system provides the first line of defense against the invasion of foreign pathogens (Mastellos et al., 2003). Its inappropriate activation may cause or aggravate several pathological conditions, including asthma, macular degeneration, rheumatoid arthritis, and rejection of xenotransplantation. The 13-residue compstatin is a promising candidate for the therapeutic treatment of unregulated complement activation (Janssen et

al., 2007; Morikis & Lambris, 2005). The conformation of the human C3 – compstatin complex is shown in Figure 4. Compstatin interacts closely with four protein sectors indicated by thick yellow tubes. An important conclusion, drawn from experimental studies with a large number of species, was that compstatin inhibits the key complement component protein C3 from several primate species, but is inactive against the corresponding protein from lower mammals, precluding the testing of compstatin-based analogues in animal models (Sahu et al., 2003). To understand this species specificity, Tamamis and co-workers compared the stabilities of compstatin complexes with human or rat C3 (Tamamis et al., 2010) by atomistic MD simulations and an MM-GBSA analysis. In the simulations of the rat C3 complex, specific protein sectors near the compstatin binding site underwent reproducible localized displacements, which eliminated or weakened critical protein-ligand interactions. In agreement with the simulations and the lack of compstatin activity against rat C3, a MM-GBSA analysis estimated the binding free energy of the human C3 complex to be stronger by -9 kcal/mol relative to the rat C3 complex, in the “single-trajectory” approximation. If protein and ligand relaxation were taken into account by a “three-trajectory” approximation, the relative binding free energy increased (in absolute value) to -19 kcal/mol. Thus, in this system the neglect of relaxation effects introduced a significant error, even though the qualitative conclusion was correct in both cases.

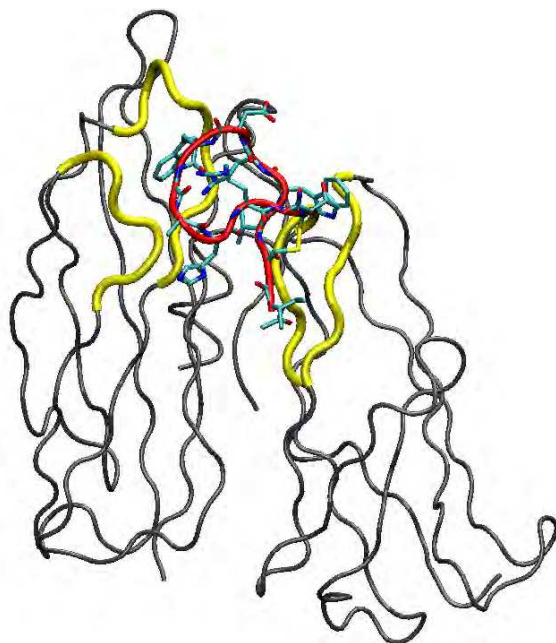


Fig. 4. Representation of the human C3c-compstatin complex. The compstatin main chain is shown in a red tube and licorice representation. The protein C3c is shown as a gray tube; only residues 329-534 and 607-620 are included, corresponding to the simulation system of Ref. (Tamamis et al, 2010). The yellow thick tubes show four protein sectors in proximity to compstatin (488-492, 454-462, 344-349 and 388-393, from left to right). The right-most sector 388-393 moves away from compstatin in the simulations of non-primate C3 complexes (Tamamis et al, 2010; Tamamis et al, 2011).

Guided by the above MM-GBSA study, subsequent simulations investigated the stabilities of compstatin complexes with “transgenic” variants of mouse C3, containing 6-9 substitutions from the human C3 sequence near the compstatin binding site. The MM-GBSA binding affinities of the resulting transgenic complexes were comparable to the one of the human C3 complex, and by -8 to -9 kcal/mol stronger relative to the mouse C3:compstatin complex (Tamamis et al., 2011). More recent simulations have investigated the affinities of a large number of human and mouse or rat C3 complexes with compstatin analogues, producing promising results (Tamamis et al., 2012). Thus, the combined study of a series of related protein-ligand complexes by atomistic simulations and an efficient evaluation of the corresponding affinities, such as in the MM-GB(PB)SA formulation, provides a powerful way to design new proteins or ligands.

4.3 Example 3: Fast predictions of binding free energies using MM-GB(PB)SA

Rastelli and co-workers (Rastelli et al., 2010) explored the reliability of using a single energy minimized receptor-ligand complex in MM-GB(PB)SA calculations to estimate ligand binding affinities for a series of structurally diverse inhibitors (Figure 5) of *Plasmodium falciparum* DHFR (Figure 6) with known binding modes and affinities.

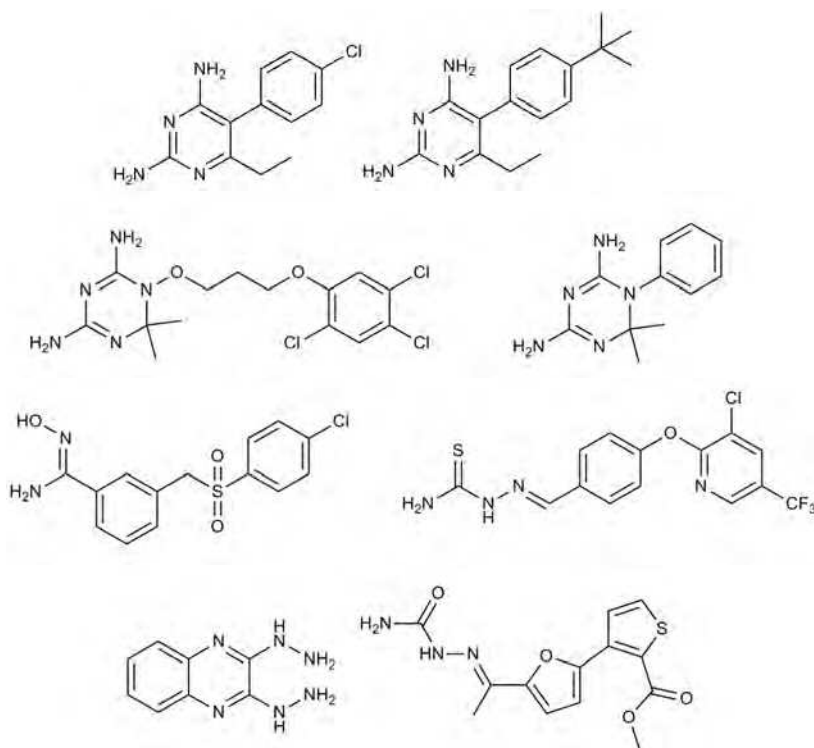


Fig. 5. Sample of the structurally diverse *Pf*DHFR inhibitors studied in Ref. (Rastelli et al., 2010).

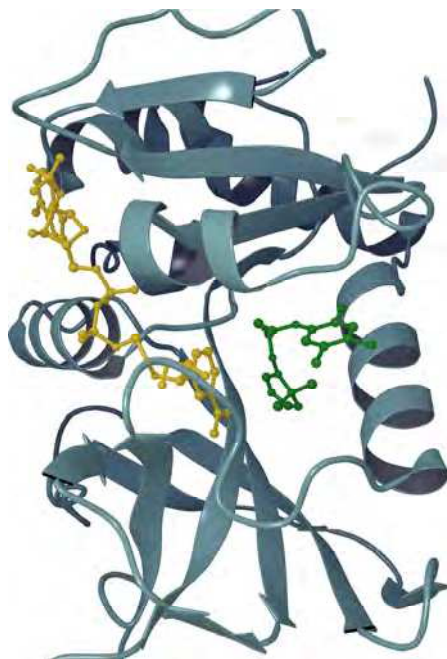


Fig. 6. Wild-type *Pf*DHFR in complex with NADPH (yellow) and the inhibitor WR92210 (green). This protein structure taken from PDB code 1J3I was used for the calculations in Ref. (Rastelli et al., 2010) and as described in Example 3.

They obtained excellent correlations between MM-PBSA or MM-GBSA binding affinities and experimental values, similar to those obtained after averaging over multiple snapshots from periodic boundary MD simulations in explicit water in the traditional sense, but with significant savings on computational time and effort. Different methods were used for generating the structures for the MM-GB(PB)SA calculations from minimizations in implicit and explicit solvent models, to minimization using a distance dependent dielectric function, and finally minimization followed by a short MD simulation and then re-minimization. The approach has been implemented in an automated workflow called BEAR (Binding Estimation After Refinement) which produces both MM-GBSA and MM-PBSA predictions of binding free energies, and is fast enough to be suitable for virtual screening applications (Degliesposti et al., 2011; Rastelli et al., 2009).

5. Conclusion

MM-GBSA and MM-PBSA are computationally efficient, end-point free energy methods that have been widely used to study protein-ligand binding affinities. Even though they lack the sound theoretical foundations of recently developed computationally demanding absolute-affinity free-energy methods (Boresch et al., 2003; Deng & Roux, 2009; Gilson & Zhou, 2007; Lee & Olsen, 2006), their connection with statistical thermodynamics has been established (Swanson et al., 2004). Due to the approximations inherent in MM-GB(PB)SA methods, they are more applicable for ranking (“scoring”) of ligand binding affinities rather

than to quantitatively predicted absolute binding free energies. They should be regarded as approximate, as they combine a molecular mechanics energy function with a continuum-electrostatics treatment of solvation effects; they include solute conformational entropy effects in an approximate manner (Singh & Warshel, 2010); and ignore the solvent molecular structure. Accurate incorporation of solute entropy (Foloppe & Hubbard, 2006) and solvent effects in binding affinity calculations is challenging, but future extensions and development of MM-GB(PB)SA methods will undoubtedly serve to address these limitations.

6. Acknowledgment

JMH acknowledges funding from the SP4-Capacities Coordination and Support Action, Support Actions, EUROSTRUCT project (CSA-SA_FP7-REGPOT-2008-1 Grant Agreement N° 230146).

7. References

- Aleksandrov, A.; Thompson, D. & Simonson, T. (2010). Alchemical free energy simulations for biological complexes: powerful but temperamental. *Journal of Molecular Recognition*, Vol.23, No.2, pp. 117-127, ISSN0952-3499
- Åqvist, J.; Luzhkov, V. B. & Brandsdal, B. O. (2002). Ligand binding affinities from MD simulations. *Accounts of Chemical Research*, Vol.35, No.6, pp. 358-365, ISSN 0001-4842
- Archontis, G.; Simonson T & Karplus M. (2001). Binding free energies and free energy components from molecular dynamics and Poisson-Boltzmann calculations. Application to amino acid recognition by aspartyl-tRNA synthetase. *Journal of Molecular Biology*, Vol.306, No.2, pp. 307-327, ISSN 0022-2836
- Archontis, G. & Simonson, T. (2005). Proton Binding to Proteins: A Free-Energy Component Analysis Using a Dielectric Continuum Model. *Biophysical Journal*, Vol.88, No.6, pp. 3888-3904, ISSN 0006-3495
- Bashford, D.; Case, D. A. (2000). Generalized Born models of macromolecular solvation effects. *Annual Review of Physical Chemistry*, 2000, Vol.51, pp. 129-152, ISSN 0066-426X
- Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F & Hermans, J. (1981). Interaction models of water in relation to protein hydration. In: *Intermolecular forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry*, B. Pullman, Editor, pp. 331-342, ISBN 902771326X, Reidel Publ. Company, Dordrecht, Holland.
- Berendsen, H. J. C.; Grigera, J. A. & Straatsma, T. P. (1987). The missing term in effective pair potentials. *Journal of Physical Chemistry*, Vol.91, No.24, pp. 6269-5271, ISSN 0022-3654
- Boresch, S.; Tettinger, F.; Leitgeb, M. & Karplus, M. (2003). Absolute binding free energies: A quantitative approach for their calculation. *Journal of Physical Chemistry B*, Vol.107, No.35, pp. 9535-9551, ISSN 1520-6106
- Boresch, S. & Bruckner, S. (2011). Avoiding the van der Waals endpoint problem using serial atomic insertion. *Journal of Computational Chemistry*, Vol.32, No.11, pp. 2449-2458, ISSN 0192-8651
- Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y. & Shaw, D. E. (2006). Scalable algorithms for molecular dynamics simulations on commodity

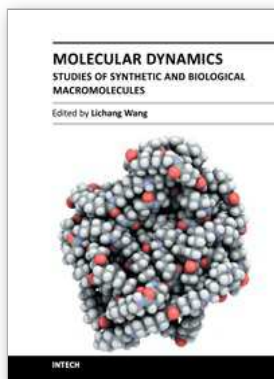
- clusters, *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, ISBN 0-7695-2700-0, Tampa, Florida, USA, Nov 11-17, 2006
- Brooks, B.R.; Janezic, D. & Karplus M. (1995). Harmonic-analysis of large systems. 1. Methodology. *Journal of Computational Chemistry*, Vol.16, No.12, pp. 1522-1542, ISSN 0192-8651
- Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B. & Woods, R. (2005). *Journal of Computational Chemistry*, Vol.26, No.16, pp. 1668-1688, ISSN 0192-8651
- Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W., Branson K. & Pande V.S. (2011). Alchemical free energy methods for drug discovery: progress and challenges. *Current Opinion in Structural Biology*, Vol.21, No.2, pp. 150-160, ISSN 0959-440X
- Cornell, W.D.; Cieplak, P.; Bayly C. I.; Gould I. R.; Merz K. M.; Ferguson D. M.; Spellmeyer D. C.; Fox T.; Caldwell J. W. & Kollman P. A. (1995). A second generation forcefield for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, Vol.117, No.19, pp. 5179-5197, ISSN 0002-7863
- Cramer, C. J. & Truhlar, D. G. (1999). Implicit solvation models: Equilibria, structure, spectra and dynamics. *Chemical Reviews*, Vol.99, No.8, pp. 2161-2200. ISSN 0009-2665
- Degliesposti, G.; Portioli, C.; Parenti, M. D. & Rastelli, G. (2011). BEAR, a novel virtual screening methodology for drug discovery. *Journal of Biomolecular Screening*, Vol.16, No.1, pp. 129-133, ISSN 1087-0571
- Deng, Y. & Roux, B. (2009). Computations of standard binding free energies with molecular dynamics simulations. *Journal of Physical Chemistry B*, Vol.113, No.8, pp. 2234-2246, ISSN 1520-6106
- Du, J.; Sun, H.; Xi, L.; Li, J.; Yang, Y.; Liu, H. & Yao, X. (2011). Molecular modeling study of Checkpoint Kinase 1 inhibitors by multiple docking strategies and Prime/MM-GBSA. *Journal of Computational Chemistry*, Vol.32, No.13, pp. 2800-2808, ISSN 0192-8651
- Dunitz, J. D. (1994). The entropic cost of bound waters in crystals and biomolecules. *Science*, Vol.264, No.5159, pp. 670-670. ISSN 0036-8075
- Foloppe, N. & Hubbard, R. (2006). Towards predictive ligand design with free-energy based computational methods. *Current Medicinal Chemistry*, Vol.13, No.29, pp. 3583-3608, ISSN 0929-8673
- Genheden, S. & Ryde, U. (2010). How to obtain statistically converged MM/GBSA results. *Journal of Computational Chemistry*, Vol.31, No.4, pp. 837-846, ISSN 0192-8651
- Gilson, M. K.; Given, J. A.; Bush, B. L. & McCammon, J. A. (1997). The statistical-thermodynamical basis for computation of binding affinities: A critical review. *Biophysical Journal*, Vol.72, No.3, pp. 1047-1069, ISSN 0006-3495
- Gilson, M. K. & Zhou, H-X. (2007). Calculation of protein-ligand binding affinities. *Annual Review of Biophysics and Biomolecular Structure*, Vol.36, pp. 21-42, ISSN 1056-8700
- Gohlke, H.; Kiel, C. & Case, D. A. (2003). Insight into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *Journal of Molecular Biology*, Vol.330, No.4, pp. 891-913, ISSN 0022-2836
- Gohlke, H. & Case, D. (2004). Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein-Protein Complex Ras-Raf. *Journal of Computational Chemistry* Vol.25, No.2, pp. 238-250, ISSN 0192-8651

- Gräter, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S. & Smith, J. C. (2005). Protein/ligand binding free energies calculated with quantum mechanics/molecular mechanics. *Journal of Physical Chemistry B*, Vol.109, No.20, pp. 10474-10483, ISSN 1520-6106
- Harvey, M. J.; Giupponi, G. & De Fabritiis, G. (2009). ACEMD: Accelerating biomolecular simulations in the microsecond time scale. *Journal of Chemical Theory & Computation*, Vol.5, No.6, pp. 1632-1639, ISSN 1549-9618.
- Hayes, J. M. & Leonidas, D. D. (2010). Computation as a tool for glycogen phosphorylase inhibitor design. *Mini Reviews in Medicinal Chemistry*, Vol.10, No.12, pp. 1156-1174, ISSN 1156-1174
- Hayes, J. M.; Skamnaki, V. T.; Archontis, G.; Lamprakis, C.; Sarrou, J.; Bischler, N.; Skaltsounis, A. L.; Zographos, S. E. & Oikonomakos, N. G. (2011). Kinetics, in silico docking, molecular dynamics, and MM-GBSA binding studies on prototype indirubins, KT5720, and staurosporine as phosphorylase kinase ATP-binding site inhibitors: The role of water molecules examined. *Proteins: Structure Function & Bioinformatics*, Vol.79, No.3, pp. 703-719, ISSN 0887-3585
- Hendsch, Z. & Tidor, B. (1999). Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Science* Vol.8, No.7, pp. 1381-1392, ISSN 0961-8368
- Hou, T.; Wang, J.; Li, Y. & Wang, W. (2011a). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information & Modeling*, Vol.51, No.1, pp. 69-82, ISSN 1549-9596
- Hou, T.; Wang, J.; Li, Y. & Wang, W. (2011b). Assessing the performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area methods. II. The accuracy of ranking poses generated from docking. *Journal of Computational Chemistry*, Vol.32, No.5, pp. 866-877, ISSN 0192-8651
- Humphrey, W.; Dalke A. & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, Vol.14, No.1, pp. 33-38, ISSN 0263-7855
- Jannsen, B. J. C.; Hallff, E. F.; Lambris, J. D. & Gros, P. (2007). Structure of compstatin in complex with complement component C3c reveals a new mechanism of complement inhibition. *Journal of Biological Chemistry*, Vol.282, No.40, pp. 29241-29247, ISSN 0021-9258
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, Vol.79, No.2, pp. 926-935, ISSN 0021-9606
- Karplus, M. & Kushick, J. N. (1981). Method for estimating the configurational entropy of macromolecules. *Macromolecules*, Vol.14, No.2, pp. 325-332, ISSN 0024-9297
- Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *Journal of Chemical Physics*, Vol.3, No.5, pp. 300-313, ISSN 0021-9606
- Kollman, P. (1993). Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews*, Vol.93, No.7, pp. 2395-2417, ISSN 0009-2665
- Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A. & Cheatham, III, T. E. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research*, Vol.33, No.12, pp. 889-897, ISSN 0001-4842

- Kongsted, J. & Ryde, U. (2009). An improved method to predict the entropy term with the MM/PBSA approach. *Journal of Computer Aided Molecular Design*, Vol.23, No.2, pp. 63-71, ISSN 0920-654X
- Kongsted, J.; Söderhjelm, P. & Ryde, U. (2009). How accurate are continuum solvation models for drug-like molecules. *Journal of Computer Aided Molecular Design*, Vol.23, No.7, pp. 395-409, ISSN 0920-654X
- Kuhn, B.; Gerber, P.; Schulz-Gasch, T. & Stahl, M. (2005). Validation and use of the MM-PBSA approach for drug discovery. *Journal of Medicinal Chemistry*, Vol.48, No.12, pp. 4040-4048, ISSN 0022-2623
- Lee, M. S. & Olson, M. A. (2006). Calculation of absolute protein-ligand binding affinity using path and endpoint approaches. *Biophysical Journal*, Vol.90, No.3, pp. 864-877, ISSN 0006-3495
- Levy, R. H., Zhang, L. Y., Gallicchio, E. & Felts, A. K. (2003). On the nonpolar hydration free energy of proteins: Surface area and continuum solvent models for the solute-solvent interaction energy. *Journal of the American Chemical Society* Vol.125, No.31, pp. 9523-9530, ISSN 0002-7863
- Lyne, P. D.; Lamb, M. L. & Saeh, J. C. (2006). Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *Journal of Medicinal Chemistry*, Vol.49, No.16, pp. 4805-4808, ISSN 0022-2623
- Lu, N. & Woolf, T. B. (2007). Chapter 6: Understanding and improving free energy calculations in molecular simulations: Error analysis and reduction methods. In: *Free Energy Calculations: Theory & Applications in Chemistry & Biology (Springer Series in Chemical Physics Vol.86)*, Chipot, C. & Pohorille, A., Springer-Verlag, ISBN 978-3-540-38447-2, Berlin-Heidelberg.
- Manta, S.; Xipnitou, A.; Kiritsis, C.; Kantsadi, A. L.; Hayes, J. M.; Skamnaki, V. T.; Lamprakis, C.; Kontou, M.; Zoumpoulakis, P.; Zographos, S. E.; Leonidas D. D. & Komiotis (2012). Contrary to docking and SAR analysis, a 3'-axial CH₂OH substitution on glucopyranose does not increase glycogen phosphorylase inhibitory potency. QM/MM-PBSA calculations suggest why. *Chemical Biology & Drug Design*, in press.
- Mahoney, M. W. & Jorgensen, W. L. (2000). A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *Journal of Chemical Physics*, Vol.112, No.20, pp.8910-8922, ISSN 0021-9606
- MacKerell, A.D.; Bashford, D.; Bellott M.; Dunbrack, Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhorn, B.; Reiher, III, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe M.; Wiórkiewicz-Kuczera, J.; Yin, D. & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, Vol.102, No.18, pp. 3586-3616, ISSN 1089-5647
- Mastellos, D.; Morikis, D.; Isaacs, S. N.; Holland, M. C.; Strey, C. W. & Lambris, J. D. (2003). Complement – Structure, functions, evolution and viral molecular mimicry. *Immunologic Research*, Vol.27, No.2-3, pp. 367-385, ISSN: 0257-277X.
- Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K. & Dill, K. A. (2007). Predicting absolute binding free energies to a simple model site. *Journal of Molecular Biology*, Vol.371, No.4, pp. 1118-1134, ISSN 0022-2836
- Morikis, D. & Lambris, J. D. (2005). Structure, dynamics, activity and function of compstatin and design of more potent analogs. In: *Structural biology of the complement system*,

- Morikis D, Lambris J. D. Editors, pp. 317-340, CRC Press/Taylor & Francis Group, ISBN 0824725409, Boca Raton, FL
- Page, C. S. & Bates, P. A. (2006). Can MM-PBSA Calculations Predict the Specificities of Protein Kinase Inhibitors? *Journal of Computational Chemistry*, Vol.27, No.16, pp.1990-2007, ISSN 0192-8651
- Pearlman, D. A. (2005). Evaluating the molecular mechanics Poisson-Boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase. *Journal of Medicinal Chemistry*, Vol.48, No.24, pp. 7796-7807, ISSN 0022-2623
- Polydoridis, S.; Leonidas, D. D.; Oikonomakos, N. G. & Archontis G. (2007). Recognition of ribonuclease A by 3'-5'-pyrophosphate-linked dinucleotide inhibitors: A molecular dynamics/continuum electrostatics analysis. *Biophysical Journal*, Vol.92, No.5, pp. 1659-1672, ISSN 0006-3495
- Rastelli, G.; Degliesposti, G.; Del Rio, A. & Sgobba, A. (2009). Binding estimation after refinement, a new automated procedure for the refinement and scoring of docked ligands in virtual screening. *Chemical Biology & Drug Design*, Vol.73, No.3, pp. 283-286, ISSN 1747-0277
- Rastelli, G.; Del Rio, A.; Degliesposti, G. & Sgobba M. (2010). Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry*, Vol.31, No.4, pp. 797-810, ISSN 0192-8651
- Rocchia, W.; Alexov, E. & Honig, B. (2001). Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *Journal of Physical Chemistry B*, Vol.105, No.28, pp. 6507-6514, ISSN 1089-5647
- Roux, B. & Simonson, T. (1999). Implicit solvent models. *Biophysical Chemistry*, Vol.78, No.1-2, pp. 1-20, ISSN 0301-4622
- Roux, B. (2001). Implicit solvent models. In: *Computational Biochemistry and Biophysics*, Becker, O.M., Mackerell, A. D. Jr, Roux, B., Watanabe, M. Editors, pp. 133-151, Marcel Dekker Inc., ISBN 0-8247-0455-X, New York
- Sahu, A.; Morikis, D. & Lambris, J. D. (2003). Compstatin, a peptide inhibitor of complement, exhibits species-specific binding to complement component C3. *Molecular Immunology*, Vol.39, No.10, pp. 557-556, ISSN 0161-5890.
- Schutz, C. N. & Warshel, A. (2001). What are the dielectric constants of proteins and how to validate electrostatic models? *Proteins-Structure, Function and Genetics*, Vol.44, No.4, pp. 400-417, ISSN 0887-3585
- Simonson, T. (2003). Electrostatics and dynamics of proteins. *Reports of Progress in Physics*, Vol.66, No.5, pp. 737-787, ISSN 0034-4885
- Singh, N. & Warshel, A. (2010). Absolute binding free energy calculations: On the accuracy of computational scoring of protein-ligand interactions. *Proteins: Structure, Function & Bioinformatics*, Vol.78, No.7, pp.1705-1723, ISSN 0887-3585
- Still, W.; C.; Tempczyk, A.; Hawley, R. & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society*, Vol.112, No.16, pp. 6127-6129, ISSN 0002-7863
- Stone, J. E.; Hardy, D. J.; Israilewitz, B. & Schulten, K. (2011). Chapter 16: GPU algorithms for molecular modeling. In: *Scientific Computing with Multicore & Accelerators*, Dongarra J.; Bader J.A. & Kurzak J., pp. (351-371), Chapman & Hall/CRC Press, ISBN 9781439825365

- Straatsma, T. P. & McCammon J. A. (1992). Computational alchemy. *Annual Review of Physical Chemistry*, Vol. 43, pp. 407-435, ISSN 0066-426X
- Swanson, J. M.; Henchman, R. H. & McCammon, J. A. (2004). Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophysics Journal*, Vol.86, No.1, pp. 67-74, ISSN 0006-3495
- Tamamis, P.; Morikis, D.; Floudas, C. A. & Archontis, G. (2010). Species specificity of the complement inhibitor compstatin investigated by all-atom molecular dynamics simulations. *Proteins: Structure, Function & Bioinformatics*, Vol.78, No 12, pp. 2655-2667, ISSN 0887-3585
- Tamamis, P; Pierou, P.; Mytidou, C.; Floudas, C. A.; Morikis, D. & Archontis, G. (2011). Design of a modified mouse protein with ligand binding properties of its human analog by molecular dynamics simulations: The case of C3 inhibition by compstatin. *Proteins: Structure, Function & Bioinformatics*, Vol. 79, No. 12, pp.3166-3179, ISSN 0887-3585
- Tamamis, P; de Victoria, A. L.; Gorham, R. D.; Bellows-Peterson, M. L.; Pierou, P.; Floudas, C. A.; Morikis, D. & Archontis G. (2012). Molecular dynamics in drug design: New generations of Compstatin analogs. *Chemical Biology & Drug Design*, in press, ISSN 1747-0285
- Thompson, D.; Plateau, P. & Simonson, T. (2006).Free-energy simulations and experiments reveal long-range electrostatic interactions and substrate-assisted specificity in an aminoacyl-tRNAsynthetase.*Chembiochem*. Vol.7, No.2, pp. 337-344, ISSN 1439-4227
- Tembe, B. L. & McCammon, J. A. (1984). Ligand-receptor interactions. *Computers & Chemistry*, Vol.8, No.4, pp. 281-283, ISSN 0097-8485
- Tidor, B. & Karplus, M. (1994). The contribution of vibrational entropy to molecular association – the dimerization of insulin. *Journal of Molecular Biology*, Vol.238, No.3, pp. 405-414, ISSN 0022-2836
- Van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krueger, P.; Mark, A. E.; Scott, W. R. P. & Tironi, I. G. (1996). Biomolecular Simulation: The GROMOS96 Manual and User Guide; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v. Zürich, Groningen
- Wang, M. & Wong, C. F. (2007). Rank-ordering protein-ligand binding affinity by a quantum mechanics/molecular mechanics/Poisson Boltzmann-surface area model. *Journal of Chemical Physics*, Vol.126, No.2, 026101, ISSN 0021-9606
- Yang, C-Y.; Sun, H.; Chen, J.; Nikolovska-Coleska, Z. & Wang, S. (2009). Importance of ligand reorganization free energy in protein-ligand binding affinity prediction. *Journal of the American Chemical Society*, Vol.131, No.38, pp. 13709-13721, ISSN 0002-7863
- Zhou, H-X. & Gilson, M. K. (2009).Theory of free energy and entropy in noncovalent binding. *Chemical Reviews*, Vol.109, No.9, pp. 4092-4107, ISSN 0009-2665
- Zhou, H-X. & Gilson, M. K. (2009).Theory of free energy and entropy in noncovalent binding. *Chemical Reviews*, Vol.109, No.9, pp. 4092-4107, ISSN 0009-2665
- Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I. Nonpolar gases. *Journal of Chemical Physics*, Vol.22, No.8, pp. 1420-1426, ISSN 0021-9606



Molecular Dynamics - Studies of Synthetic and Biological Macromolecules

Edited by Prof. Lichang Wang

ISBN 978-953-51-0444-5

Hard cover, 432 pages

Publisher InTech

Published online 11, April, 2012

Published in print edition April, 2012

Molecular Dynamics is a two-volume compendium of the ever-growing applications of molecular dynamics simulations to solve a wider range of scientific and engineering challenges. The contents illustrate the rapid progress on molecular dynamics simulations in many fields of science and technology, such as nanotechnology, energy research, and biology, due to the advances of new dynamics theories and the extraordinary power of today's computers. This second book begins with an introduction of molecular dynamics simulations to macromolecules and then illustrates the computer experiments using molecular dynamics simulations in the studies of synthetic and biological macromolecules, plasmas, and nanomachines. Coverage of this book includes: Complex formation and dynamics of polymers Dynamics of lipid bilayers, peptides, DNA, RNA, and proteins Complex liquids and plasmas Dynamics of molecules on surfaces Nanofluidics and nanomachines

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Joseph M. Hayes and Georgios Archontis (2012). MM-GB(PB)SA Calculations of Protein-Ligand Binding Free Energies, *Molecular Dynamics - Studies of Synthetic and Biological Macromolecules*, Prof. Lichang Wang (Ed.), ISBN: 978-953-51-0444-5, InTech, Available from: <http://www.intechopen.com/books/molecular-dynamics-studies-of-synthetic-and-biological-macromolecules/mm-gb-pb-sa-calculations-of-protein-ligand-binding-free-energies>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.