Can MM/GBSA calculations be sped up by system truncation?

Misini Ignjatović, Majda; Mikulskis, Paulius; Söderhjelm, Pär; Ryde, Ulf

Published in:
Journal of Computational Chemistry

DOI:
10.1002/jcc.25120

2018

Document Version:
Early version, also known as pre-print

Can MM/GBSA calculations be sped up by system truncation?

Majda Misini Ignjatović [a], Paulius Mikulskis [a], Pär Söderhjelm [b], Ulf Ryde [a,*]

[a] Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden

[b] Department of Biophysical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden

Correspondence to Ulf Ryde, E-mail: Ulf.Ryde@teokem.lu.se, Tel: +46 – 46 2224502, Fax: +46 – 46 2228648

2017-11-07

Abstract

We have studied whether calculations of the binding free energy of small ligands to a protein by the MM/GBSA approach (molecular mechanics combined with generalised Born and surface area solvation) can be sped up by including only a restricted number of atoms close to the ligand. If the protein is truncated before the molecular dynamics (MD) simulations, quite large changes are observed for the calculated binding energies, e.g. 4 kJ/mol average difference for a radius of 19 Å for the binding of nine phenol derivatives to ferritin. The results are improved if no atoms are fixed in the simulations, with average and maximum errors of 2 and 3 kJ/mol at 19 Å and 3 and 6 kJ/mol at 7 Å. Similar results are obtained for two additional proteins, p38α MAP kinase and factor Xa. On the other hand, if energies are calculated on snapshots that are truncated after the MD simulation, all residues more than 8.5 Å from the ligand can be omitted without changing the energies by more than 1 kJ/mol on average (maximum error 1.4 kJ/mol). At the molecular mechanics level, the gain in computer time for such an approach is small. However, it shows what size of system should be used if the energies instead are calculated with a more demanding method, e.g. quantum-mechanics.

Key Words: Ligand-binding affinities, MM/GBSA, system truncation, ferritin, p38α MAP kinase, factor Xa.

Introduction

An important goal in computational chemistry is to develop methods to estimate the binding affinity of a small molecule (L) to a biological macromolecule (R), i.e. the free energy (ΔG_{bind}) of the reaction

$$R + L \leftrightarrow RL \quad (1)$$

If ΔG_{bind} could be calculated accurately, significant parts of drug development could be performed computationally. This would reduce the need of synthesis and experimental testing of thousands of putative drug candidates, thereby decreasing the time consumption and cost of the drug development.

Consequently, many methods have been developed with this aim [1,2], ranging from simple scoring and docking methods [3,4], which can screen thousands of candidates in a short time, to strict free-energy perturbation (FEP) approaches [5,6], which in principle should give the correct results if the sampling is exhaustive and the energy function is perfect, but at a large computational cost. The so-called end-point approaches are intermediate in computational effort and statistical-mechanics rigour. They are based on a physics-based force field and conformational sampling with molecular dynamics (MD) simulations, but only of the three reactants in Eqn. 1. Typical examples are the PDLD/s-LRA/β (semi-macroscopic protein-dipoles Langevin-dipoles method within a linear-response approximation) [7], linear-interaction energy (LIE) [8,9], MM/PBSA, and MM/GBSA (molecular mechanics combined with Poisson–Boltzmann or generalised Born and surface-area solvation) [10,11,12], and they have been employed for many different systems with a varying success [12,13,14,15,16].

Still, such calculations are rather expensive, requiring 10–100 ns MD simulation for each RL complex. As the computational effort of MD simulations scales with the square of the number of atoms in the simulated system (n log n with Ewald summation), a natural attempt to reduce the effort is to make the simulated system smaller. System truncations in MD simulations have a long history [17,18,19,20,21,22]. In the early days of MD, it was often necessary to make the calculations possible, but as the computers became more powerful, the need and interest of truncations decreased. However, some groups still employ truncated systems, in particular for LIE calculations [23]. Recently, we showed that relative binding free energies calculated by FEP did not change by more than 1.4 kJ/mol if the size of the simulated system was reduced from 46 to 15 Å (from 38844 to 1480 atoms) [24].

The question then naturally arises whether MM/GBSA calculations can be sped up in a similar manner by system truncations. In fact, many early MM/PBSA studies employed systems truncated at a radius of ~20 Å [11,25]. In one study, MM/PBSA results obtained with a truncated system were compared with those obtained with the complete protein, simulated both in a spherical system with reaction-field corrections and in a periodic box, treated with Ewald summation [26]. All three approaches gave results of a similar quality compared to experimental data, but the statistical precision of the method was too poor to decide whether the various approaches gave significantly different results.

In this paper, we perform a systematic study of the effect of system truncations on binding affinities obtained with the MM/GBSA approach. Starting from simulations of a full system, we successively remove peripheral parts of the protein and examine how the calculated affinities change. As test cases, we employ the binding of nine phenol derivatives to a ferritin dimer [27], 17 ligands to p38α MAP kinase (p38) [28], and ten ligands to blood-clotting factor Xa (fXa) [29]. These systems have been used in several previous computational studies of binding affinities [24,29,30,31,32,33,34,35,36,37].

Methods

The studied systems

We have studied ferritin with nine phenol-derivatives with alkane substituents at the ortho positions (L01–L09) [27], p38α MAP kinase with 17 ligands (P1–P17) [28], and blood clotting factor Xa with ten ligands [29]. The ligands are shown in Figure 1. The preparation of the proteins and the ligands has been described before [24,31,36]. Functional ferritin is a multimer of almost 100 protein chains, but following earlier investigations, it was modelled as a dimer (the ligand binds at the interface between two subunits) [27,30,31,32]. All ionisable protein residues were assigned their standard protonation state at pH 7 and the protonation of histidine residues was determined from the solvent-accessibility and hydrogen-bond interactions. In each subunit of ferritin, His-49, 132, and 147 were modelled as doubly protonated, His-114 was protonated on the ND1 atom, and His-124 on the NE2 atom. In p38, His-107 and 313 were assumed to be protonated on ND1, His-174 on both atoms, and the other nine His residues only on the NE2 atom. In fXa, His-57 and 83 were protonated on ND1, His-13 on both N atoms, and the other three His residues on the NE2 atom. The Amber99SB force field was used to describe the protein [38], whereas the ligands were described by the general Amber force field [39]. Ligand charges were calculated with the restrained electrostatic potential method [40] using potentials calculated at the HF/6-31* level and sampled with the Merz–Kollman scheme [41]. Each protein–ligand system was immersed in an octahedral box of TIP3P [42] water molecules that extended at least 10 Å outside the protein.

MD simulations

The MD simulations were run by the sander module in Amber 11 [43]. The temperature was kept at 300 K using Langevin dynamics [44] with a collision frequency of 2.0 ps^{-1}. The pressure was kept at 1 atm using a weak-coupling approach [45] with isotropic position rescaling and a relaxation time of 1 ps. Particle-mesh Ewald summation [46] with a fourth-order B-spline interpolation and a tolerance of 10^{-5} was used to treat long-range electrostatics and the long-range van der Waals interactions were treated with a continuum approach. The non-bonded cutoff was 8 Å and the non-bonded pair list was updated every 50 fs. The MD time step was 2 fs and the SHAKE algorithm [47] was used to constrain bond lengths involving hydrogen atoms.

The protein–ligand complexes were simulated in the following way: First, the complex was optimised by 500 steps of steepest descent minimisation, keeping all atoms, except water molecules and hydrogen atoms, restrained to their starting positions with a force constant of 418 kJ/mol/$Å^2$. The minimisation was followed by 20 ps MD equilibration with a constant pressure and the restraining force reduced to 214 kJ/mol/$Å^2$. Finally, a 1200 ps simulation was performed at a constant pressure, but without any restrains. During the last 200 ps, 40 snapshots were collected, which were used for MM/GBSA energy calculations.

MM/GBSA calculations

In the MM/GBSA approach, the free energy of each reactant is calculated from

$$G = E_{\text{int}} + E_{\text{el}} + E_{\text{vdW}} + G_{\text{pol}} + G_{\text{np}} \quad (2)$$

where the first three terms are standard molecular mechanics (MM) energies from internal (bond, angle and dihedral), electrostatic, and van der Waals interactions. They were calculated with Amber 11 [43] using the same force field as in the simulations, after stripping off all water molecules and without any periodic boundary conditions, but with an infinite cutoff. G_{pol} and G_{np} are the polar and non-polar continuum contributions to the solvation free energies. G_{pol} was calculated by the GB method by Onufriev, Bashford, and Case, model I [48]. The non-polar solvation energy was estimated from the solvent-accessible surface area (SASA) according to $G_{np} = \gamma$ SASA + b, with $\gamma = 0.0227$ (kJ/mol)/Å^2 and $b = 3.85$ kJ/mol [25]. Sometimes, an entropy term is also included in MM/GBSA, estimated by a normal-mode analysis of the vibrational frequencies calculated at the MM level. However, in this study, this term was not considered because it typically already involves truncated systems (or is omitted) [10,11,12,49] and the size dependence of this term has recently been studied [33]. It is dominated by the loss of translational and rotational entropy of the ligand, which is independent of the truncation and similar for ligands of a similar size [12,50].

The binding free energy was calculated by using Eqn. 2 for the complex, the free ligand, and the free protein:

$$\Delta G_{bind} = \langle G_{RL} - G_R - G_L \rangle_{RL} \quad (3)$$

Here, the brackets indicate an average over 40 snapshots taken from the MD simulation of the complex. The structures of the free protein and the free ligand were obtained by simply removing the other part of the complex. Thereby, the precision is improved and the E_{int} term cancels [11,12]. To obtain statistically converged results, we employed 40 independent simulations for each of the ligands, by assigning different starting velocities to atoms (i.e. 48 ns simulations and 1600 energy calculations in total for each ligand) [51]. Previous studies have shown that many short independent simulations cover the phase space more effective than a single long simulation and give a more representative estimate of the statistical uncertainty [35,51,52,53,54,55,56]. Reported uncertainties are the standard deviation over the results (averages over the 40 snapshots) from the 40 independent simulations, divided by $\sqrt{40}$.

Truncation

Truncated systems were obtained by removing all residues that do not have any atom within a certain radius (R) from any atom in the largest ligand (L02 for ferritin, P6 for p38, and ligand **5** for fXa). Truncation radii of R = 19, 14, 10, 8.5 and 7 Å were tested for ferritin and 35, 30, 25, 20, 15, and 10 Å for the other two (larger) proteins. Four different sets of truncations were performed for ferritin as is described in Table 1: truncation after the MD simulation (T1) or truncation before the MD simulation, based either on equilibrated structures (T3) or on the crystal structure. Two different types of truncations were employed for the latter simulations (T2 and T4, see below). For the T1 and T3 truncations, all 40 × 40 MD snapshots were used to define the system and any residue having at least one atom within R of L02 in any of the snapshots were retained. In contrast, the truncations of the crystal structure (T2 and T4) were based on a single structure (viz. the protonated crystal structure), but still using the L02 ligand. Therefore, these truncated systems were significantly smaller (see Table 2). All truncations with the same R (and truncation scheme) contained exactly the same protein atoms for the nine ligands.

For truncations T1–T3, non-consecutive residues were capped by acetyl and N-

methylamine groups, but no such capping was performed for T4. In the MD simulations (T2–T4), all atoms in the residues that do not have any atom within a radius of R – 4 Å from the ligand (in any of the snapshots used to define the truncation) were kept fixed at the starting coordinates (crystallographic or equilibrated structure), to avoid that atoms move in an unrealistic manner when atoms are removed. However, in one set of calculations (T4f), all atoms were free to move, to estimate the effect of the fixed atoms. Figure 2 shows the full protein and the most truncated system (R = 7 Å), based on the crystal structure (T2).

In the T2 simulations, the system was solvated in a pre-equilibrated sphere of TIP3P waters up to a radius of R + 4 Å from the geometrical centre of the system. The system was simulated without periodic boundary conditions but with a spherical restraint keeping the water molecules inside a sphere with a radius of R + 4 Å. In the T3 simulations, only water molecules within R of the ligand in the original snapshot from equilibration were kept (thus, the number of water molecules varied depending on the starting snapshot). The systems were simulated with spherical restraints on the water molecules with a radius of R Å. The T2 and T3 simulations employed a non-bonded cutoff of 16 Å. In the T4 simulations, the truncated systems were solvated in a pre-equilibrated truncated octahedral box of TIP3P waters extending at least 3, 4, 5, 6, or 8 Å from any atom in the system for the truncation radii of R = 7, 8.5, 10, 14, and 19 Å, respectively. The systems were then simulated with periodic boundary conditions and a constant pressure. These simulations treated electrostatics with particle-mesh Ewald summations and a cutoff of 8 Å for the van der Waals interactions.

For truncations T2–T4 we run the same simulations as for the full system: First, a minimisation with all atoms, except water molecules and hydrogen atoms, restrained to their starting positions with a force constant of 418 kJ/mol/Å^2, then, a 20 ps MD simulation with the restraining force reduced to 214 kJ/mol/Å^2, and finally a 1200 ps simulation without any restrains, collecting 40 snapshots for MM/GBSA energy calculations during the last 200 ps. In all simulations, residues outside R – 4 Å were fixed, except in the T4f simulations. For p38 and fXa, only T4 and T4f truncations were tested.

Result and Discussion

In this paper, we have made a systematic study of the effect of system truncations in MM/GBSA calculations of absolute affinities for nine phenol derivatives (shown in Figure 1a) binding to the ferritin dimer. Five sets of calculations were performed, depending on how and when the truncation was performed, as is described in Table 1. The size of the various systems is shown in Table 2 and in Figure 2. The results of the five truncation schemes are described in separate sections. To check that the results apply also to other proteins, we have also studied the binding of 17 ligands of p38α MAP kinase and ten ligands of blood-clotting factor Xa (shown in Figures 1b and 1c).

Truncation after the MD simulations

In the first set of calculations on ferritin, the MD simulations were run with the full (not truncated) protein. After the MD simulations, energies were calculated on snapshots from the MD simulation, in which we included only residues within a certain radius from the ligand (R = 7, 8.5, 10, 14, and 19 Å; truncation scheme T1). The results of these calculations are shown in Figure 3. It can be seen that the truncations had a small effect on the absolute MM/GBSA energies (Figure 3a): At R = 10 Å, the maximum error for all ligands was 0.9 kJ/mol (average error 0.6 kJ/mol) and even for the smallest radius (R = 7 Å), all errors were below 2.5 kJ/mol (1.7 kJ/mol on average). The relative binding affinities (relative to the average over the nine ligands) were even more stable: The errors in the relative affinities were always below 1.0

kJ/mol, with an unsigned average that increased from 0.03 kJ/mol at R = 19 Å to 0.4 kJ/mol at R = 7 Å.

The error relative to the calculations on the full system quite closely followed the size of the ligand: The smallest ligand L09 gave the smallest error and the largest ligand L02 gave the largest error. The error was always positive (i.e. the truncated calculations gave a less negative ΔG_{bind} than calculations with the full protein. This effect came mainly from the E_{vdW} term, which was always positive and increased when R is decreased (Figure 3b). This term represents the truncation of the long-range dispersion, which is always attractive. The E_{el} and G_{pol} terms always had opposite signs and therefore partly cancelled (Figure 3c), as is frequently observed in MM/GBSA calculations [12,26]. However, the error of their sum was positive with only two minimal exceptions for L09, thereby reinforcing the error in the E_{vdW} term (but the sum was typically half as large as the E_{vdW} term, up to 0.8 kJ/mol). The error in G_{np} was always negligible (<0.01 kJ/mol).

The standard errors for the truncated systems were similar or slightly larger than those for the full protein, 0.7–1.4 compared to 0.6–1.0 kJ/mol, for all nine ligands and truncation radii. Consequently, we can conclude that it is possible to truncate the snapshots after the MD simulations, with only minimal changes in the calculated affinities. This represents the intrinsic effect of the truncation, i.e. when the geometry is not changed. However, the gain in computational time is minimal, because the MD simulations completely dominate the time consumption.

Truncation of the crystal structure, before MD simulations

Next, we tried to truncate the systems before the MD simulations, so that both the simulations and energy calculations were performed on truncated systems. The natural choice would be to truncate the protonated crystal structure and then run both the MD simulation and MM/GBSA energy calculations on the truncated systems (truncation scheme T2). Unfortunately, the results in Figure 4 showed that the error of the truncation in this case is ~10 times larger than if the systems are truncated after the MD simulations, with errors of up to 7 kJ/mol already for the largest radius (R = 19 Å). The error increased up to 24 kJ/mol for the smallest systems, but the variation was much less systematic than that in Figure 3.

The precision of the MM/GBSA energies from truncated calculations was similar to that of the original calculations, with standard errors of 0.2–1.1 kJ/mol for all calculations, and with a slightly decreasing trend when R is decreased (average standard error decreasing from 0.8 to 0.5 kJ/mol). Thus, the precision was not the problem.

For most of the ligands and radii, the error compared to the full protein was still positive, but in nine cases, negative errors were found (down to –5 kJ/mol). As before, the errors in the E_{el} and G_{pol} terms had in most cases opposite signs and partly cancelled (except in three cases for L08 and L09 for R = 19 and 14 Å, when both were positive, but one of them small). However, their sum could be both positive and negative, and up to 14 kJ/mol. The error in the E_{vdW} term could also have a varying sign, although it was always positive for the two smallest radii. It typically showed an increasing trend for the smaller radii, except for ligand L07. It also typically dominated the other terms, especially for small radii. The error in the G_{np} term was still below 1 kJ/mol, with a varying sign (but typically negative, at least for the smaller radii).

Unfortunately, the relative binding affinities also showed sizeable errors. They were up to 7–17 kJ/mol for the various truncations with unsigned averages of 3–5 kJ/mol, with only a slight increase when the truncation radius is decreased.

Truncation after MD equilibration

Next, we tried to first run a single 200 ps equilibration MD simulation with full ferritin for each ligand, before the systems were truncated and restarted with new velocities for the 40 independent MD simulations (truncation scheme T3). This was an attempt to reduce possible van der Waals clashes in the crystal structure before the systems were truncated. In addition, these simulations were fully solvated inside a sphere with a radius of R.

The results of these calculations are shown in Figure 5. It can be seen that the errors were similar to those in Figure 4: The maximum error increased from 8 kJ/mol for R = 19 Å to 18 kJ/mol for the smallest radii (average errors from 3 to 13 kJ/mol). The errors were quite varying, with large differences between the ligands and somewhat unclear trends. The precision was slightly worse than for the calculations started from the crystal structure, with standard errors of 0.6–1.8 kJ/mol for all calculations, but with a decreasing trend when R was decreased (average standard error decreasing from 1.5 to 1.1 kJ/mol).

The error in the E_{vdW} term was in general positive (except for R = 19 Å and for L09). It showed an increasing trend as R was decreased in most cases. The error in E_{el} was typically negative and that in G_{pol} was always positive. Strangely, they both typically showed a decreasing trend when R decreased. As an effect, their sum was positive (except for the two largest radii with L06) and quite small (up to 7 kJ/mol) with a maximum at R = 14 Å. The error in G_{np} was always less than 1 kJ/mol and was in general positive, except sometimes for R = 7 Å.

Again, the errors in the relative binding affinities were essentially independent of the truncation radius. They were up to 7–9 kJ/mol for the various radii, with unsigned averages of 3–6 kJ/mol. Unfortunately, this is still too large to be really useful.

Truncation from crystal structure with periodic systems

Finally, we performed a fourth set of truncations (T4), based on the crystal structure of ferritin, but without using any capping groups and employing fully solvated octahedral periodic systems (as in the original simulations of the full system), extending 3–8 Å from the truncated protein. As for truncation schemes T2 and T3, residues with at least one atom within R – 4 Å of the ligand were allowed to move, whereas the other heavy protein atoms were fixed at their crystal-structure positions.

Unfortunately, the truncation errors in these calculations were 2–5 times larger than for the spherical systems (T2 and T3): A comparison of the calculations of the full and the truncated systems showed that the average absolute error over the nine ligands was 9–12 kJ/mol for R = 14–19 Å and increased up to 52 kJ/mol when R = 7 Å (Figure 6). In general, the error depended on the size of the ligand: Big ligands exhibited large errors, which tended to increase much when reducing R. Errors were lower for the smaller ligands (L07, L08, and L09) and they were less affected by R. For the relative binding affinities, the trends were somewhat more regular: The average absolute errors increased regularly from 3 kJ/mol for R = 19 Å to 41 kJ/mol for R = 7 Å.

The error in the G_{np} term was still small and positive, up to 2 kJ/mol. The error in the E_{el} term was negative with a single exception. For the larger ligands, L01–L03, L05, and L06, it was rather small (up to 12 kJ/mol) with a limited variation. For the other ligands, it was larger (up to 32 kJ/mol) and more varying. The error in the G_{pol} term was always positive. For most ligands, it increased with decreasing truncation radii, up to 8–38 kJ/mol. Therefore, the sum of the E_{el} and G_{pol} terms was in general positive, at least for the smaller radii, except for the two smallest ligands L08 and L09. The error in the E_{vdW} term for the larger ligands L01–L06 was small and negative for the two largest R, but became large and positive for the smaller

radii. Thereby, it dominated over the other terms. For the small ligands, it was smaller, so that the error in the E_{el} and G_{pol} terms also contributed to the net effect in ΔG_{bind}.

The standard errors varied for the different systems. It was small for the smaller ligands (L07, L08, and L09), 0.2–1.1 kJ/mol, but for the larger ligands (L01–L06) it was larger and it increased with decreasing R, up to 2–6 kJ/mol for R = 7 Å. The reason for this is that the ligand rotates in the active site, giving two or three distinct conformations with widely different calculated binding affinities, as is shown in Figure 7. Similar effects, but to lesser extent, were observed also in the T3 calculations.

To avoid the rotation of the large ligands, we ran additional simulations at R = 7 Å, in which residues with no atom within 2 Å (rather than 3 Å) of the ligand were kept fixed. For some ligands, this was successful: The standard error was 1 kJ/mol or less for five of the ligands. However, for L01, L03, L05, and L06, the standard error was still 3–6 kJ/mol and the calculated ΔG_{bind} showed several distinct binding modes (in fact, L07 and L08 also showed grouping of the binding affinities, although the difference is smaller, so that the standard error remains ~1 kJ/mol). Interestingly, this led to a larger deviation from the results of the full system for all nine ligands as can be seen in Figure 8. This was caused mainly by the E_{vdW} term. The effect was particularly pronounced for the small L07 and L08 ligands. This indicates that there are two opposing effects of the fixed residues: One is to avoid that the systems move away from the crystal structure. The other is to allow enough fluctuations in the system.

Consequently, we performed also a set of calculations with the same truncation scheme, but with no atoms kept fixed (T4f). Somewhat unexpectedly, this improved the results very much. As can be seen in Figure 9, the truncation error was up to 3 kJ/mol for the two largest radii and increased up to 6 kJ/mol for the smallest radius (mean absolute errors over the nine ligands of 1.3–3.0 kJ/mol). The errors were even smaller and more regular for the relative binding free energies: The mean and maximum errors increased from 1.2 and 2 kJ/mol at R = 19 Å to 2.7 and 5 kJ/mol at R = 7 Å, respectively.

The precision was also better than in the T4 calculations: The standard error increased from 0.7–1.1 kJ/mol for the largest truncation radius to 1.1–2.4 kJ/mol for the smallest radius. For four ligands with the largest standard errors (L01, L03, and L06 at R = 7 Å and L02 at R = 8.5 and 10 Å), 1–3 of the independent simulations still gave diverging results, indicating rotation of the ligand. However, these were so few that they did not affect the net results significantly.

From Figure 9b, it can be seen that the error in E_{vdW} was rather small and with a varying sign, –5 to +4 kJ/mol, although all calculations with the two largest radii gave a too positive E_{vdW}. The error in E_{el} was somewhat larger and negative, except in three cases, –13 to 5 kJ/mol. It was counteracted by the G_{GB} term (–2 to 12 kJ/mol) with a single exception. Consequently, the $E_{el} + G_{GB}$ sum was quite small (of a magnitude similar to the E_{vdW} term), with a varying sign, –4 to 5 kJ/mol. The error in the G_{np} term was always minimal, up to 0.5 kJ/mol.

Thus, we can conclude that for ferritin it is much more important to allow a full flexibility of the binding site than keeping the residues close to their positions in the crystal structure. A compromise could be to instead restraining residues with no atom within R – 4 Å of the ligand towards the starting crystal structures. We tested also such an approach (with a force constant of 41 kJ/mol/Å^2), but it gave worse results.

P38α MAP kinase and factor Xa

All calculations so far were performed on ferritin with the nine small and neutral ligands

in Figure 1a. In order to ensure that results apply also to other systems, we have tested truncated MM/GBSA calculations also on two additional systems, viz. the binding of 17 ligands of p38α MAP kinase (p38) and ten ligands of blood-clotting factor Xa (fXa) [28,29,32]. The ligands are shown in Figures 1b and 1c. It can be seen that they are larger and more drug-like than the ferritin ligands. All p38 ligands are neutral, whereas two of the fXa ligands (**39** and **63**) have a single and the others a double positive charge. Calculations were performed on the full proteins, as well with the truncation schemes T4 and T4f using truncation radii of R = 35, 30, 25, 20, 15, and 10 Å.

The results for p38 are shown in Figure 10. It can be seen that with the T4 scheme, the truncation errors are quite large for all truncation radii, with an (absolute) average increasing from 2.5 to 7.9 kJ/mol as the radius decreases from 35 to 10 Å. Likewise, the maximum error increases from 8 to 16 kJ/mol. This is somewhat better than for ferritin, but still too large to be of any practical use. If no atoms were fixed (T4f, shown in Figure 10b), the results improved significantly: The average errors decreased to 1.7–5.5 kJ/mol and the maximum errors to 4–11 kJ/mol. Relative energies are even better with average and maximum errors of 2.8 and 7 kJ/mol for the smallest radius. Again, the results are similar (slightly worse) than for ferritin.

The results for fXa are shown in Figure 11. In this case, the truncation errors for the T4 scheme are larger and therefore more similar to those of ferritin: The average error increases from 3.6 kJ/mol at the two largest radii to 12 kJ/mol at R = 15 Å and 46 kJ/mol at R = 10 Å. The maximum error is 9–61 kJ/mol. As expected, the truncation error is dominated by the electrostatic and solvation terms, which are 340–500 kJ/mol for the smallest radius for the doubly charged ligands (140–160 kJ/mol for the singly charged ligands). As usual the two terms have opposing signs, but for the smallest radius the cancellation is rather poor, giving a net effect of –41 to –57 kJ/mol for the doubly charged ligands. This is much larger than the van der Waals term, which contributes by up to 13 kJ/mol. However, for $R \geq 20$ Å, the two terms are of a similar magnitude.

Without any fixed atoms, (T4f in Figure 11 b), the errors are somewhat smaller, but the improvement is smaller than for the other test cases: The average error is 3.0–29 kJ/mol and the maximum error 7–36 Å. Again, R = 10 Å gives much worse results than R = 15 Å (4.5 and 12 kJ/mol). Again, the error is dominated by the electrostatic and GB terms, but already at R = 15 Å is their sum of the same magnitude as the van der Waals term. Again, the results are strongly improved if only relative energies are considered: Then the average errors are 2.6–4.6 kJ/mol and the maximum errors 8–11 kJ/mol.

Conclusions

In this paper, we have studied whether MM/GBSA calculations can be sped up by using truncated systems, in the same way as has been observed for FEP calculations [24]. The calculations on ferritin show that if the MD simulations were run with the full protein and only the energy calculations were performed on the truncated systems, the results were satisfactory and all atoms more than 8.5 Å from the ligand could safely be removed without changing the calculated absolute binding affinities by more than 1 kJ/mol on average. However, this led to only a minimal gain in computer time, as the MD simulations (and the entropy calculations) dominate the time consumption of the MM/GBSA calculations.

Unfortunately, if the systems were truncated before the MD simulations, large changes in the calculated energies were observed in the calculated energies, e.g. an average error of 4 kJ/mol for ferritin even for the largest radius. Running first a MD equilibration on a solvated full protein and then truncating before the production MD simulations were run

improved the results slightly, but the average error was still 3 kJ/mol for R = 19 Å and increased to 13 kJ/mol for the smallest radii. The relative affinities showed errors that were virtually independent of the truncation radius, but still quite sizeable, 3–6 kJ/mol on average for both truncations. If we instead used fully solvated octahedral systems for the truncations, the errors were even larger, 3–41 kJ/mol even for the relative binding affinities. In addition, rotation of the ligands was observed, giving rise to distinct groups of calculated binding affinities. This indicates that system truncation is not accurate enough to be used to speed up MM/GBSA calculations. However, the problem could be traced to the constraint of some atoms during the simulations: If the atoms were not fixed, the average errors (in both absolute and relative binding free energies) were reduced to 1–3 kJ/mol for all the radii. However, the maximum errors were still quite sizeable, up to 6 kJ/mol for the smallest radii.

Thus, the results show that truncation is much more problematic than for FEP calculations. In particular, we have shown that fixing atoms is problematic. Atoms at the periphery of the simulated system are traditionally fixed in simulations of truncated systems to avoid that they move in a way that would not be possible if the entire protein was simulated. However, this also introduces the risk that the system becomes too rigid. Our results show that the latter effect is much more important. Unfortunately, it cannot be known beforehand which of the two effects is dominating.

Consequently, we have to conclude that the effect of fixing atoms has to be tested for each binding site. Moreover, our results indicate that the truncation errors are quite sizeable, 2 kJ/mol on average even for R = 19 Å and 3 kJ/mol for R = 7 Å. The user has to decide whether the gain in computation time justifies errors of this size. Table 3 shows the time consumption for the various truncated simulations. It can be seen that T2 gives the largest speed-up, because these calculations give the smallest simulated systems. For R = 7 Å, the truncated simulations are 76 times faster than for the full system, a large saving in computer time. The T4 truncations are 2–4 times slower, because they involve much larger systems (owing to the added water molecules in the periodic box). The difference increases when R is decreased, but it is smaller than what could be expected from the difference in the number of atoms, because simulations with periodic systems are faster as they can use a smaller cutoff radius (owing to the Ewald treatment of the long-range electrostatics). This is also the reason why the T3 calculations are slowest, 3–6 times slower than T2 (increasing when R is decreased) and 1.3–1.7 times slower than T2.

Similar calculations on two other systems, 17 neutral ligands of p38α MAP kinase and ten positively charged ligands of factor Xa show similar results: The truncation errors are quite large, but are to some extent improved by not keeping atoms fixed. For charged ligands, the errors are larger, especially at truncation radii below 15 Å. This indicates that the results may be rather general.

Therefore, we believe that the present results have stronger bearing on the use of quantum-mechanical methods to calculate binding affinities [57]. They could be used as a post-processing of MM/GBSA calculations, to reduce the effect of inaccuracies in the MM force field. In fact, truncation schemes T1–T3 were selected with such an aim (e.g. by using capped amino acids). In such calculations, only minimal truncation errors are acceptable, but on the other hand the energy calculations will dominate the time consumption. Undoubtedly, the conclusion would then be that the MD simulations should be performed with the full protein, but the quantum-mechanical calculations can then be performed on truncated systems with a radius of down to 8.5 Å without introducing a truncation error of more than 1 kJ/mol on average.

Acknowledgements

This investigation has been supported by grants from the Swedish research council (projects 2010-5025 and 2014-5540) and from Knut and Alice Wallenberg Foundation (KAW 2013.0022). The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University and HPC2N at Umeå University.

References

1 H. Gohlke, G. Klebe *Angew. Chem. Int.* **2002**, *41*, 2644-2676.
2 K. A. Sharp in Protein–ligand interactions, H Gohlke, ed. Wiley-VCH Verlag GmbH & Co, 2012, pp 3-22.
3 R. Rajmani, A. C. Good *Curr. Opin. Drug. Discov. Develop.* **2007**, *10*, 308-315.
4 C. Sotriffer in Protein–ligand interactions, H Gohlke, ed. Wiley-VCH Verlag GmbH & Co, 2012, pp 237-263
5 J. Michel, J. W. Essex JW *J. Comput.-Aided Mol. Design* **2010**, *24*, 639-658
6 T. Steinbrecher in Protein–ligand interactions, H Gohlke, ed. Wiley-VCH Verlag GmbH & Co, 2012, pp 207-236
7 Y. Y. Sham, Chu ZT, Tao H, Warshel A. *Proteins* **2000**, *39*, 393–407
8 J. Åqvist, Medina C, Samuelsson J-E. *Prot. Eng.* **1994**, *7*, 385–391
9 T. Hansson, Marelius J, Åqvist J. J Comput-Aided Mol Des **1998**,12, 27–35
10 J. Srinivasan, Cheatham TE, Cieplak P, Kollman PA, Case DA. *J Am Chem Soc* **1998**, 120, 9401–9409
11 P. A. Kollman, Massova I, Reyes C, Kuh, B., Huo, S, Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, Pl, Srinivasa, J., Case, D. A., Cheatham, T. E. Acc Chem Res **2000**, *33*, 889–897
12 S. Genheden, U. Ryde (**2015**) Expert Opinion Drug Discov., *10*, 449-461
13 B. O. Brandsdal, Österberg, F., Almlöf, M., Feierberg, I., Luzhkov, V. B., Åqvist J. Adv. Prot. Chem. **2003**, *66*, 123-158.
14 N. Foloppe, Hubbard R. Curr Med Chem **2006**, *13*, 3583-3608
15 T. Hou, Wang J, Li Y, Wang W. J Chem Inf Model **2011**, *51*, 69-82
16 P. A. Greenidge, Kramer C, Mozziconacci J-C, Wolf RM. J Chem Inf Model **2012**, *53*, 201-209
17 M. Berkowitz, McCammon, J. A. *Chem. Phys. Lett.* **1982**, *90*, 215-217.
18 A. T. Brünger, Brooks III, C. L., Karplus, M. *Chem. Phys. Lett.* **1984**, *105*, 495-500.
19 G. King, Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647–366
20 J. W. Essex, Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 951-972.
21 W. Im, Bernèche, S., Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924-2937.
22 N. K. Banavali, Im, W., Roux, B. *J. Chem. Phys.* **2002**, *117*, 7381-7388.
23 J. Marelius, Kolmodin, K., Feierberg, I., Åqvist, J. *J Mol. Graph. Model.* **1998**, *16*, 213-225.
24 S. Genheden, Ryde, U. J. Chem. Theory Comput., **2012**, *8*, 1449-1458
25 B. Kuhn, Kollman, P. A. J. Med. Chem. **2001**, *43*, 3786-3791.
26 A. Weis, Katebzadeh K, Söderhjelm P, Nilsson I, Ryde U (**2006**) J Med Chem *49*, 6596-6606
27 L. S. Vedula, Brannigan, G., Economou, N. J., Xi, J., Hall, M. A., Liu, R., Rossi, M. J., Dailey, W. P., Grasty, K. C., Klein, M. L., Eckenhoff, R. G., Loll, P. J. A Unitary Anesthetic Binding Site at High Resolution. *J. Biol. Chem.* **2009**, *284*, 24176-24184

28 D. A. Pearlman, Charifson, P. S. *J. Med. Chem.*, **2001**, *44*, 3417–3423.
29 H. Matter, Defossa, E., Heinelt, U., Blohm, P. M., Schneider, D., Müller, A., Herok, S., Schreuder, H., Lie, A., Brachvogel, V., Lönze, P., Walser, A., Al-Obeidi, F., Wildgoose, P. *J. Med. Chem.*, **2002**, *45*, 2749-2769.
30 P. Mikulskis, Genheden, S., Wichmann, K., Ryde, U. (**2012**) *J. Comput. Chem*. *33*, 1179-1189.
31 S. Genheden, P. Mikulskis, L. Hu, J. Kongsted, P. Söderhjelm, U. Ryde, *J. Am. Chem. Soc*., **2011**, *133*, 13081-13092.
32 S. Genheden, P. Mikulskis, U. Ryde (**2014**), J. Chem. Inf. Model., *54*, 2794-2806
33 Genheden, S., Kuhn, O., Mikulskis, P., Hoffmann, D., U. Ryde *J. Chem. Inf. Model*. , **2012**, *52*, 2079-2088
34 P. Mikulskis, S. Genheden, U. Ryde (**2014**) *J. Mol. Model*., *20*, 2273.
35 S. Genheden, U. Ryde *J. Comput. Chem.* **2011**, *32*, 187
36 S. Genheden, I. Nilsson, U. Ryde *J. Chem. Inf. Model*., **2011**, *51*, 947-958
37 Genheden, S., Ryde, U. *Proteins*, **2012**, *80*, 1326-1342
38 V. Hornak, Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. *Proteins, Struct., Funct. Bioinform*. **2006**, *65*, 712
39 J. M. Wang, Wolf, R. M., Caldwell, K. W., Kollman, P. A., Case, D. A. *J. Comput. Chem*. **2004**, *25*, 1157
40 C. I. Bayly, Cieplak, P., Cornell, W. D., Kollman, P. A. *J. Phys. Chem*. **1993**, *97*, 10269
41 B. H. Besler, Merz, K. M., Kollman, P. A. *J. Comput. Chem*. **1990**, *11*, 431
42 W. L. Jorgensen, Chandrasekhar, J., Madura, J. D., Impley, R. W., Klein, M. L. *J Chem Phys* **1983**, *79*, 926
43 D. A. Case, Darden, T.A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Lou, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, B., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L, Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, G. G., Seetin, M. G., Sagui, C., Babin, V., Kollman, P. A. AMBER 10, University of California, San Francisco, 2008.
44 X. Wu, Brooks, B. R. *Chem. Phys. Lett*. **2003**, *381*, 512
45 H. J. C. Berendsen, Postman, J. P. M., van Gunsteren, W. F., DiNola, A., Haak, J R. *J. Chem. Phys*. **1984**, 81, 3684
46 T. Darden, York, D., Pedersen, L. *J. Chem. Phys*. **1993**, *98*, 10089
47 J. P. Ryckaert, Ciccotti, G., Berendsen, H. J. C. *J. Comput. Phys*. **1977**, *23*, 327
48 A. Onufriev, Bashford, D., Case, D. A. *Proteins*, **2004**, *55*, 383-394.
49 J. Kongsted, Ryde, U. *J. Comp.-Aided Mol. Des*. **2009**, *23*, 63
50 J. Kongsted, U. Ryde (**2009**) *J. Comput. Aided Mol Design*, *23*, 63.
51 S. Genheden, U. Ryde (**2010**) *J. Comput. Chem.*, *31*, 837-846
52 A. Elofsson, Nilsson, L. *J Mol Biol* **1993**, *233*, 766.
53 M. Lawrenz, R. Baron, J. A. McCammon *J. Chem. Theory Comput.* **2009**, *5*, 1106.
54 B. Zagrovic, van Gunsteren, W. F. J Chem *Theory Comput* 2007, 3, 301.
55 M. Adler, P. Beroza *J. Chem. Inf. Model.* **2013**, *53*, 2065.
56 H. J. Wijma, S. J. Marrink, D. B. Janssen, *J. Chem. Inf. Model.* **2014**, *54*, 2079.
57 U. Ryde, P. Söderhjelm *Chem. Rev*. **2016**, *116*, 5520-5566

Table 1. Description of the various truncation schemes. In T1, no MD simulations were performed, in contrast to the other three schemes. The starting structure was either the crystal structure, the final structure after a short MD equilibration or after the full MD simulation. The residues are either capped with acetyl and N-methylamide groups or just truncated between the amino acids. In some MD simulations, all atoms in residues that do not have any atom within a radius of $R - 4$ Å from the ligand were kept fixed at the starting coordinates. The simulated systems were either periodic (truncated octahedron), spherical (non-periodic with a restraint to keep atoms within the sphere), or just the non-periodic system obtained after truncation. Water molecules were always taken from the truncation (those within the radius R of the ligand) but additional water molecules were sometimes added from a pre-equilibrated water box, giving a spherical or truncated octahedral system.

Name	MD	Starting structure	Capped residues	Fixed atoms	Simulated system	Solvation
T1	no	after full MD	yes	no MD	truncated	from truncation
T2	yes	crystal	yes	yes	spherical	added R + 4 Å sphere
T3	yes	after short MD equilibration	yes	yes	spherical	from truncation
T4	yes	crystal	no	yes	periodic	added R + 3–8 Å box
T4f	yes	crystal	no	no	periodic	added R + 3–8 Å box

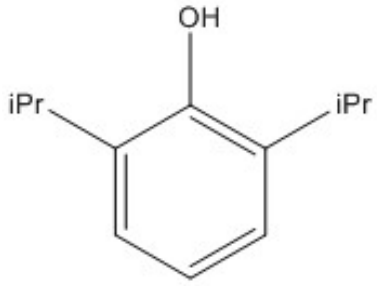
Table 2. Number of protein atoms (atm), moving protein atoms (mov), amino acids (res), capping groups (cap), and water molecules (water) in the various truncated calculations (T1–T4) for ferritin.

	T1 and T3					T2					T4			
R	atm	mov	res	caps	water	atm	mov	res	caps	water	atm	mov	res	water
∞	5432	5432	340	0	16268–9	5432	5432	340	0	16268–9	5432	5432	340	16268–16269
19	4643	3253	275	12	582–747	3819	2423	220	16	451–480	3669	1963	222	4933–4940
14	3348	2010	186	22	168–249	2462	1377	131	20	87–109	2344	966	139	2953–2955
10	2285	1196	117	15	34–58	1509	587	83	9	8–14	1474	320	87	1311–1316
8.5	1786	866	97	10	17–31	1284	290	71	8	2–8	1144	175	67	758–768
7	1433	467	80	6	9–15	984	207	54	8	1–5	880	73	53	515–519

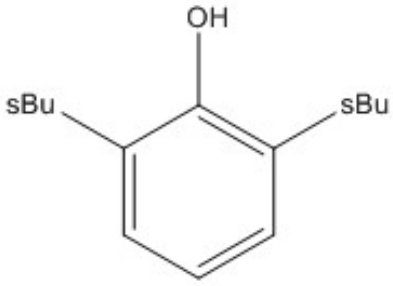
Table 3. Approximate timing for the various simulations, presented as speed-up relative to the full simulations ($R = \infty$; 1 indicates no speed-up).

	R =					
	∞	19	14	10	8.5	7
T1	1	1	1	1	1	1
T2	1	5.1	13	30	51	76
T3	1	1.7	3.2	7.6	10	13
T4	1	3.0	4.8	10	15	22

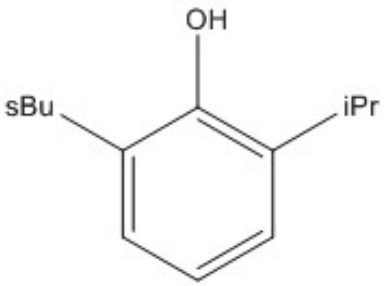
Figure 1. (a) The nine phenol derivatives considered in this study. Et, Pr, iPr, and sBu denote ethyl, propyl, isopropyl and secondary butyl groups, respectively. (b) The 17 ligands of p38 [28]. The inserted text shows the substituents in the order (R_1, R_2, R_3). (c) The structure of the ten fXa ligands, named as in the original study [29]. The reference compound **53** has $R_1 = R_2 = R_3 = R_5 = H$, $R_4 = NH$, $R_6 = –C(NH_2)_2^+$. The other ligands differ from **53** in the following ways: **5**: R_1 = Br, R_2 = Me; **9**: R_1 = Cl; **39**: R_2 = OH, $R_5 = R_6$ = OMe; **47**: R_2 = Me; **49**: $R_3 = NH_2$; **50**: R_2 = OH; **51**: R_3 = OH; **63**: R_2 = OH, R_6 = OMe; **125**: R_4 = O.



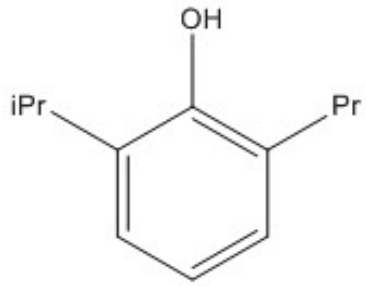
L01



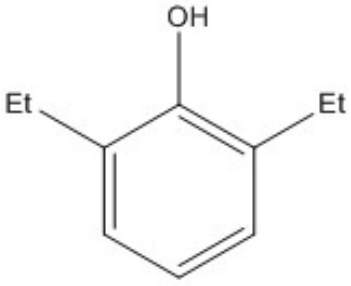
L02



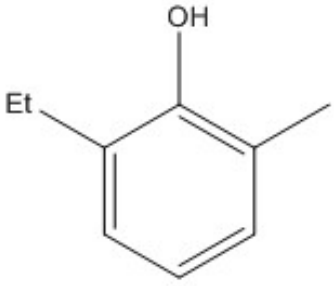
L03



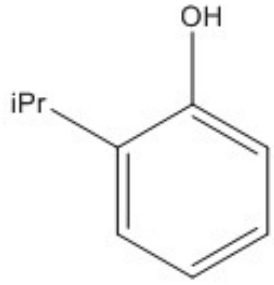
L04



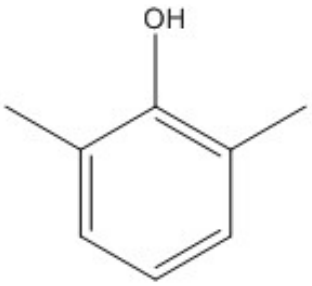
L05



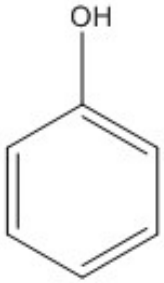
L06



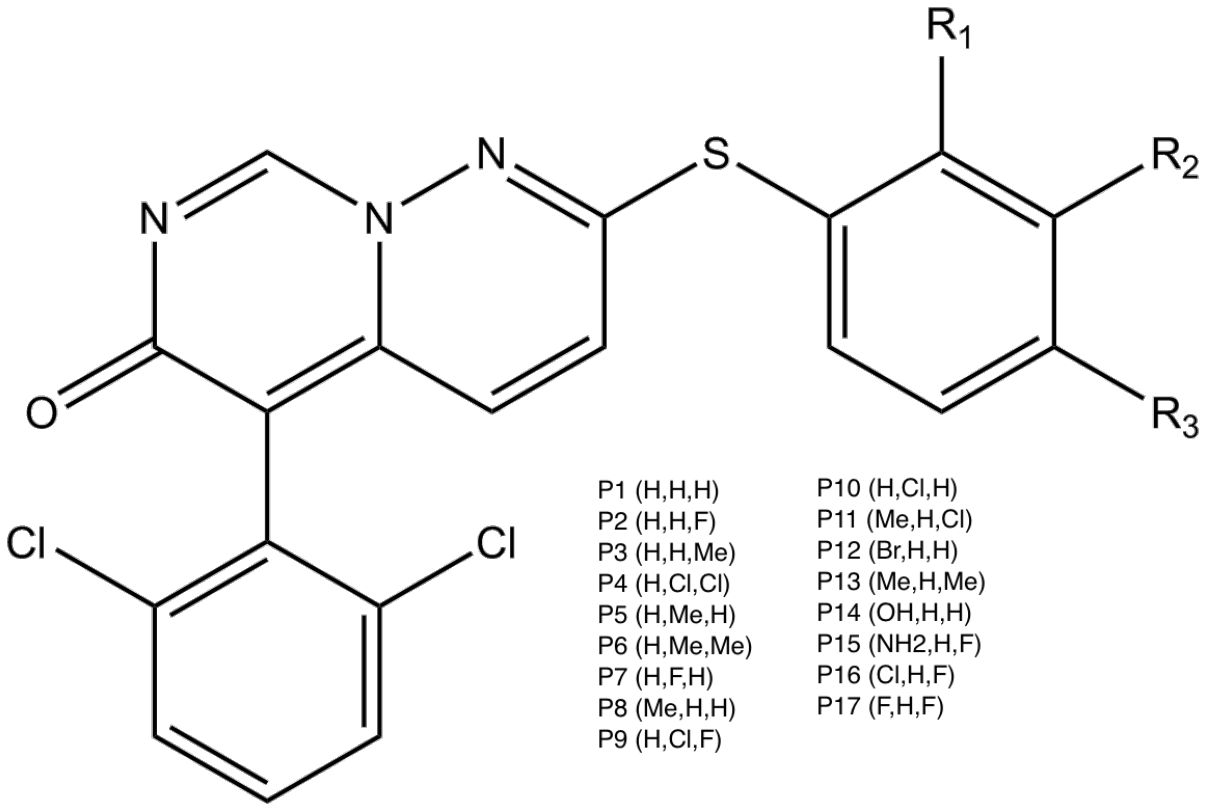
L07



L08



L09



P1 (H,H,H)
P2 (H,H,F)
P3 (H,H,Me)
P4 (H,Cl,Cl)
P5 (H,Me,H)
P6 (H,Me,Me)
P7 (H,F,H)
P8 (Me,H,H)
P9 (H,Cl,F)
P10 (H,Cl,H)
P11 (Me,H,Cl)
P12 (Br,H,H)
P13 (Me,H,Me)
P14 (OH,H,H)
P15 (NH2,H,F)
P16 (Cl,H,F)
P17 (F,H,F)

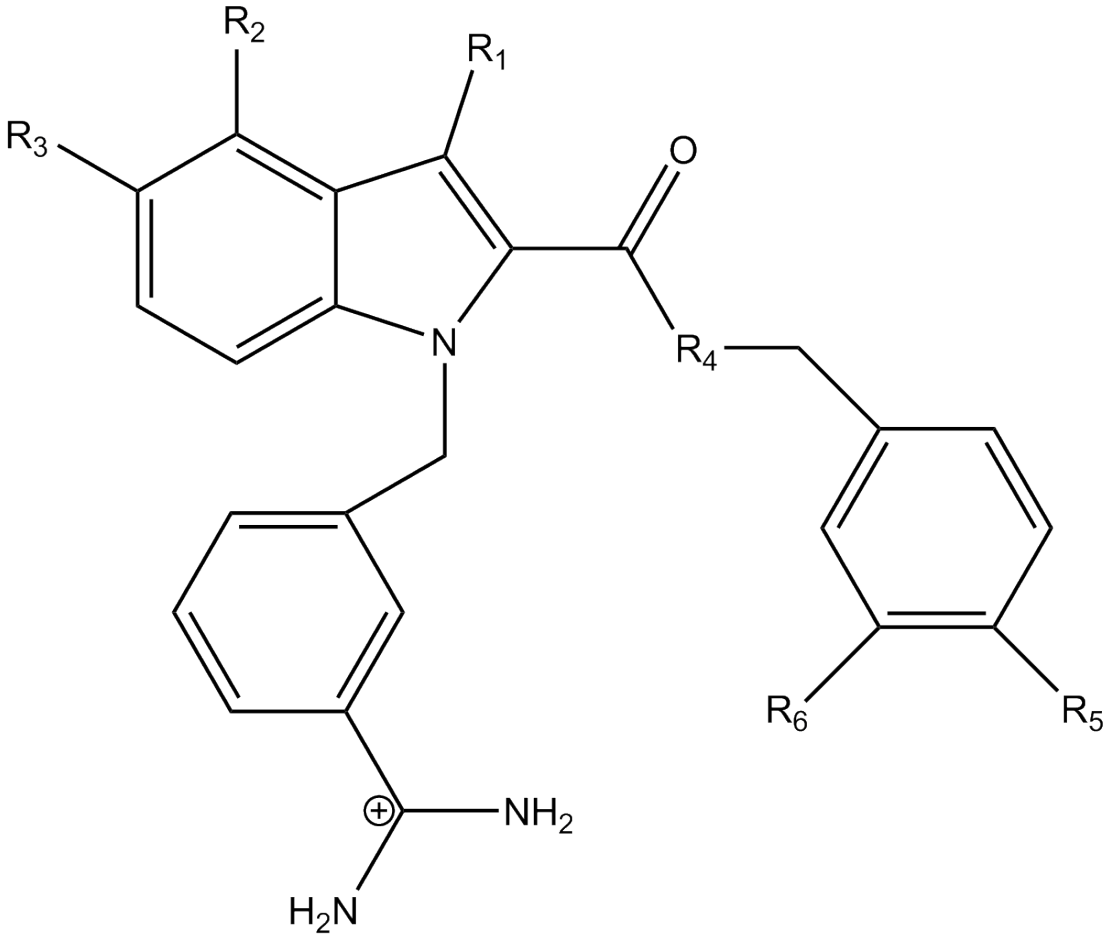


Figure 2. The full simulated ferritin dimer (left) and the most truncated system (R = 7 Å), based on the crystal structure (T2; right). In both figures, the ligand (L02) is shown in green with a space-filling model, whereas fixed atoms are shown with thin tubes and moving atoms with thick tubes (only right model). Water molecules are omitted for clarity.

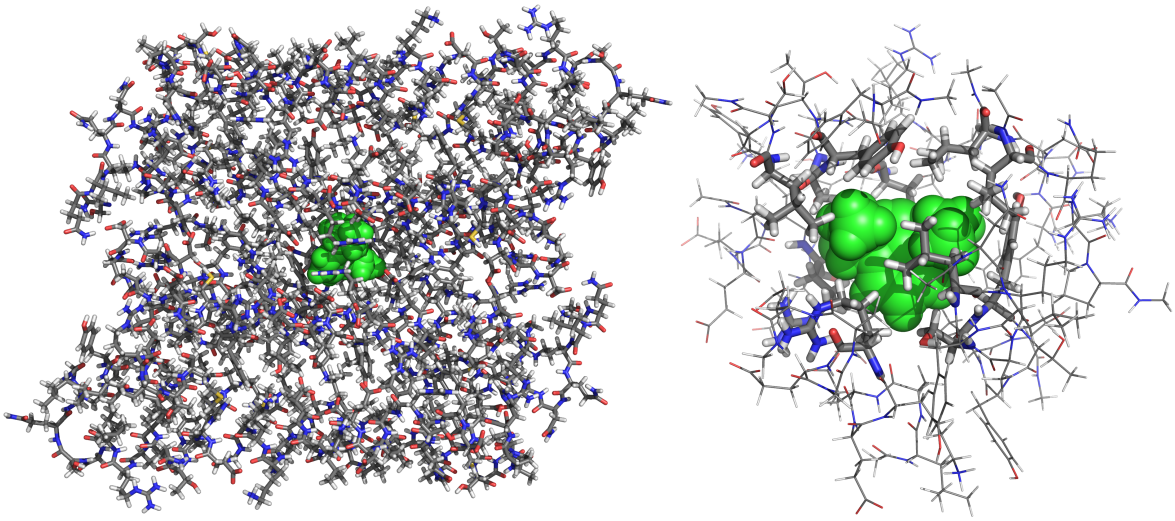
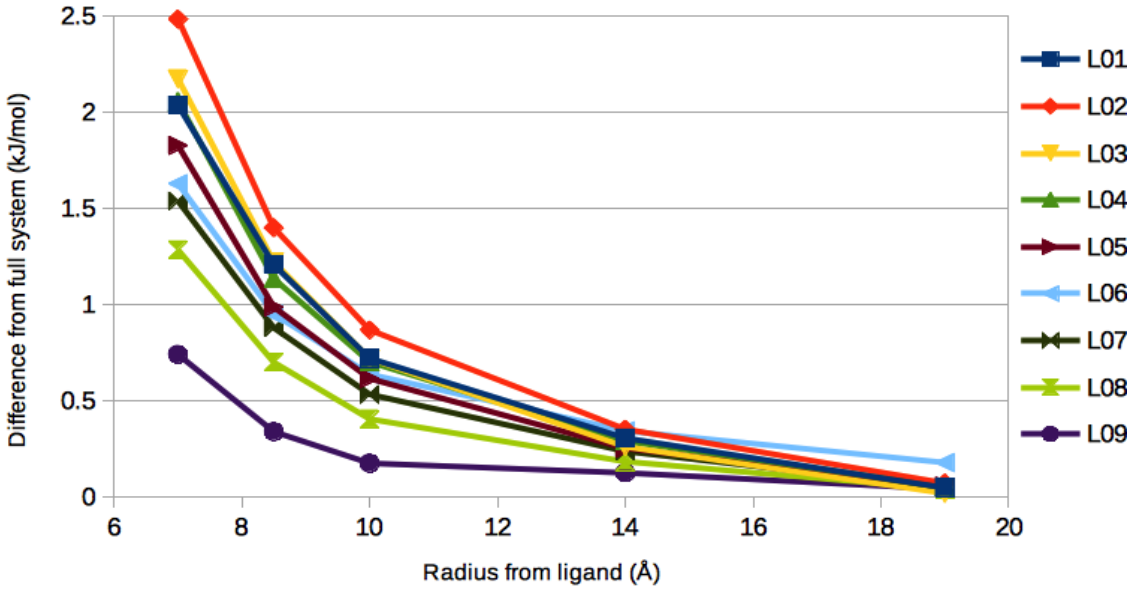
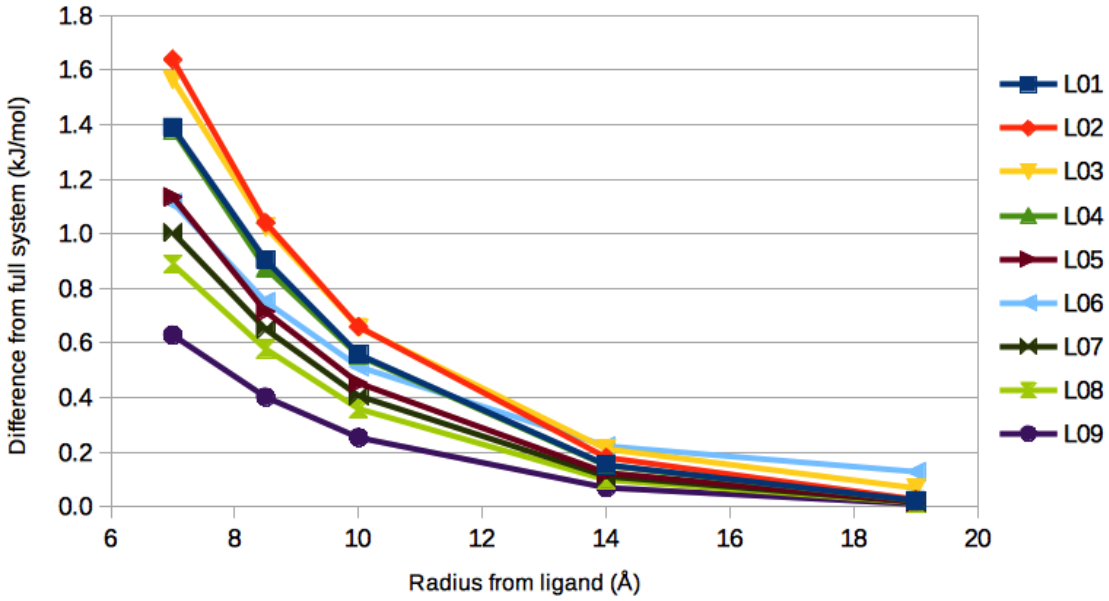


Figure 3. Effect of system truncations after the MD simulation (T1; i.e. truncation only for the MM/GBSA energy calculations) for ferritin. The difference in total MM/GBSA absolute binding energy between the calculations employing the full and the truncated systems is shown (kJ/mol) for each ligand as a function of the truncation radius. The top figure shows total energies, the middle figure E_{vdW}, and the lower figure E_{el}, G_{GB}, and $E_{el} + G_{GB}$ (full, dotted, and dashed lines, respectively). G_{np} is always < 0.01 kJ/mol and is therefore not shown.





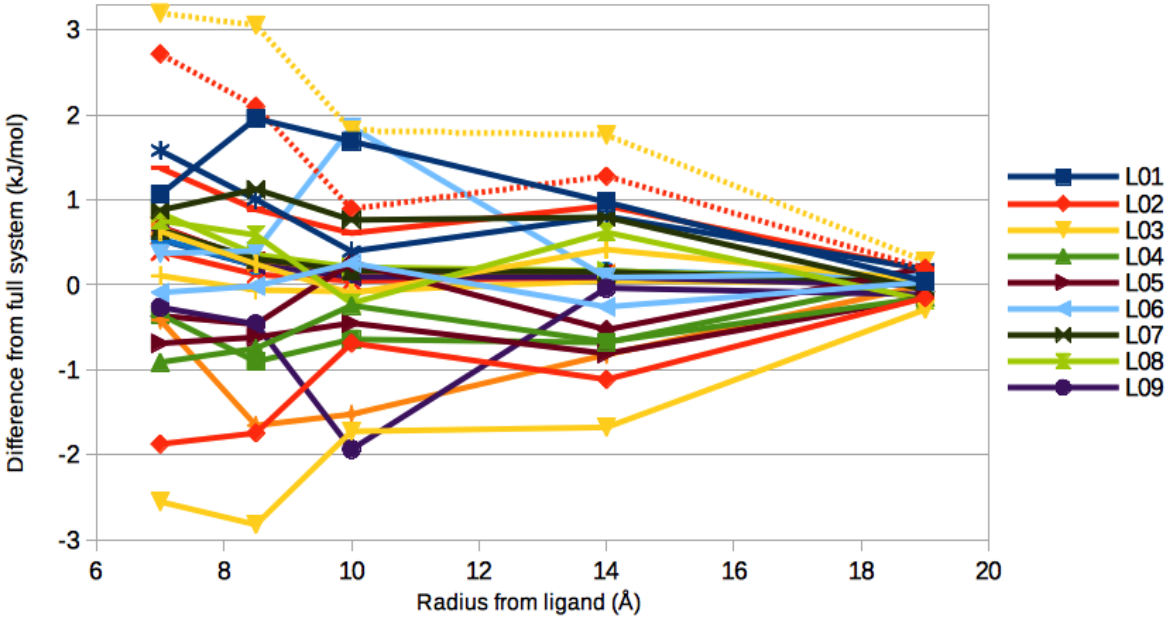
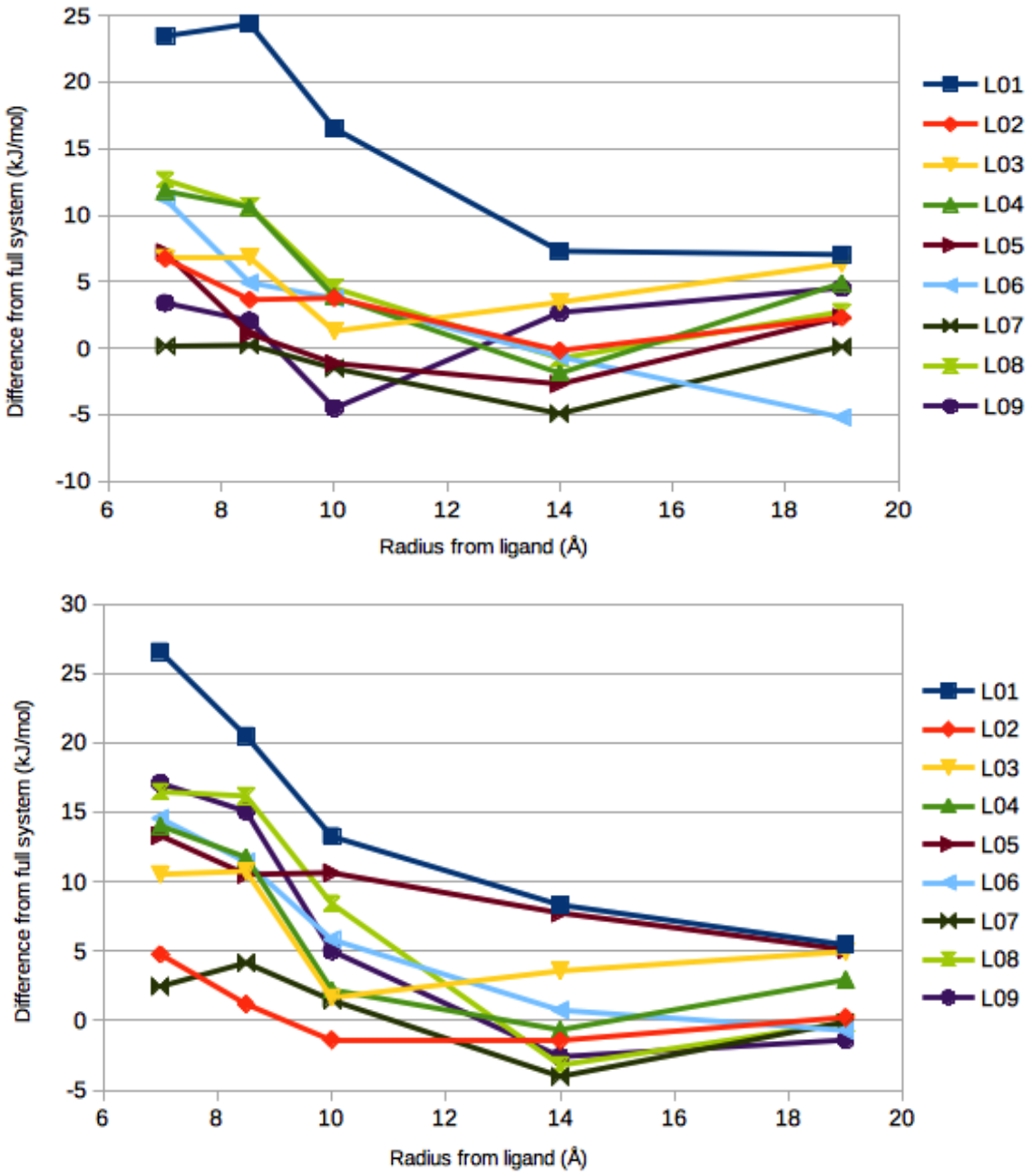

Difference from full system (kJ/mol)
3
2
1
0
-1
-2
-3
6
8
10
12
14
16
18
20
Radius from ligand (Å)
L01
L02
L03
L04
L05
L06
L07
L08
L09

Figure 4. Effect of system truncations for MM/GBSA calculations on ferritin based on MD simulations run on truncated systems started from the solvated crystal structure (T2). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius. The top figure shows total energies, the middle figure E_{vdW}, and the lower figure E_{el}, G_{GB}, and $E_{el} + G_{GB}$ (full, dotted, and dashed lines, respectively). G_{np} is always between –1 and 1 kJ/mol and is not shown.



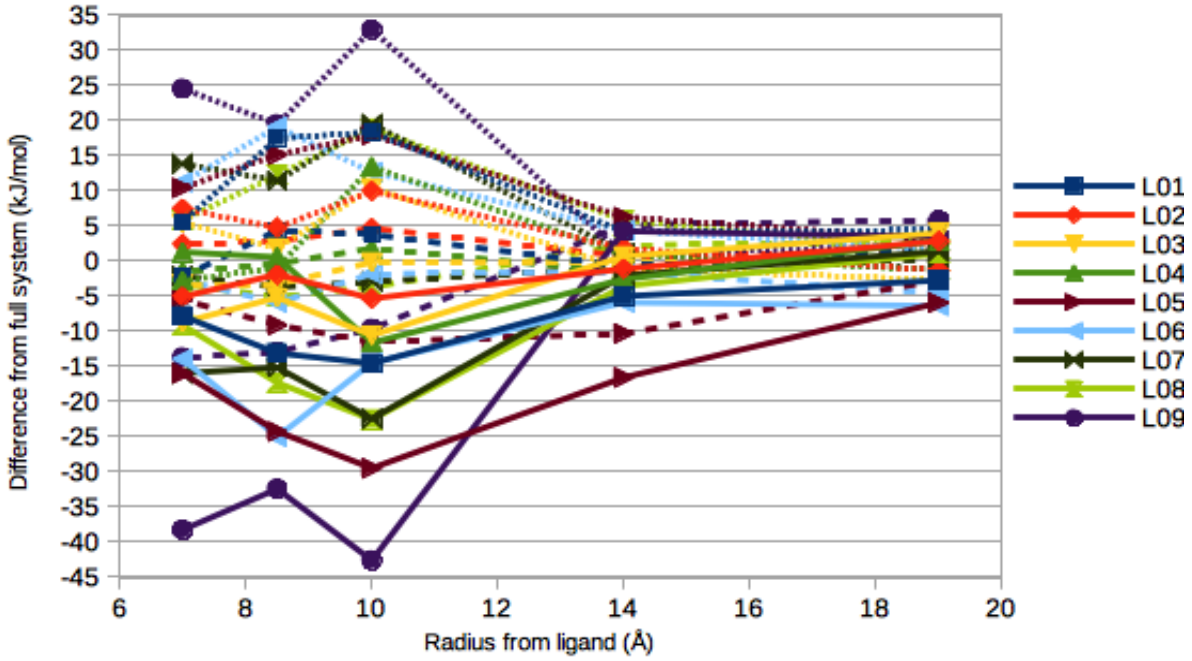
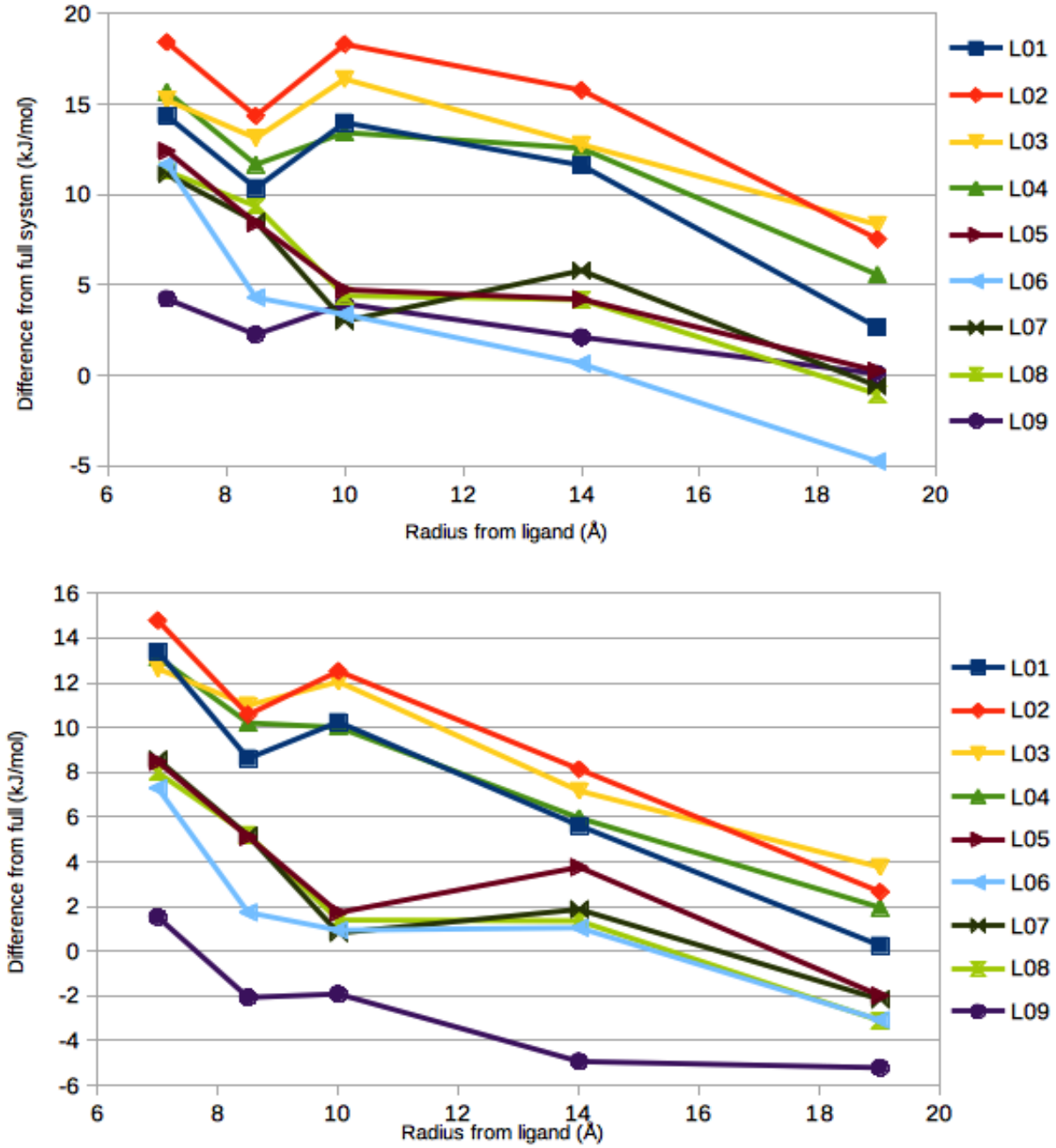

Difference from full system (kJ/mol)
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
6
8
10
12
14
16
18
20
Radius from ligand (Å)
L01
L02
L03
L04
L05
L06
L07
L08
L09

Figure 5. Effect of system truncations for MM/GBSA calculations based on MD simulations run on truncated systems started from a structure, equilibrated with the full protein (T3). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius. The top figure shows total energies, the middle figure E_{vdW}, and the lower figure E_{el}, G_{GB}, and $E_{el} + G_{GB}$ (full, dotted, and dashed lines, respectively). G_{np} is always between 0 and 1 kJ/mol and is not shown.



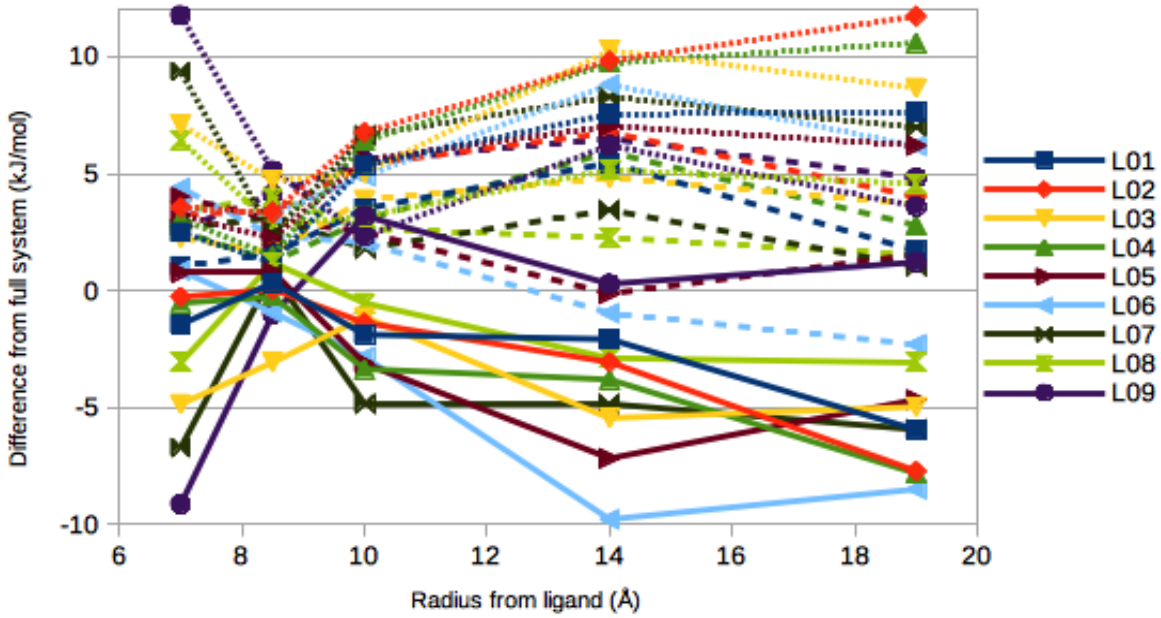
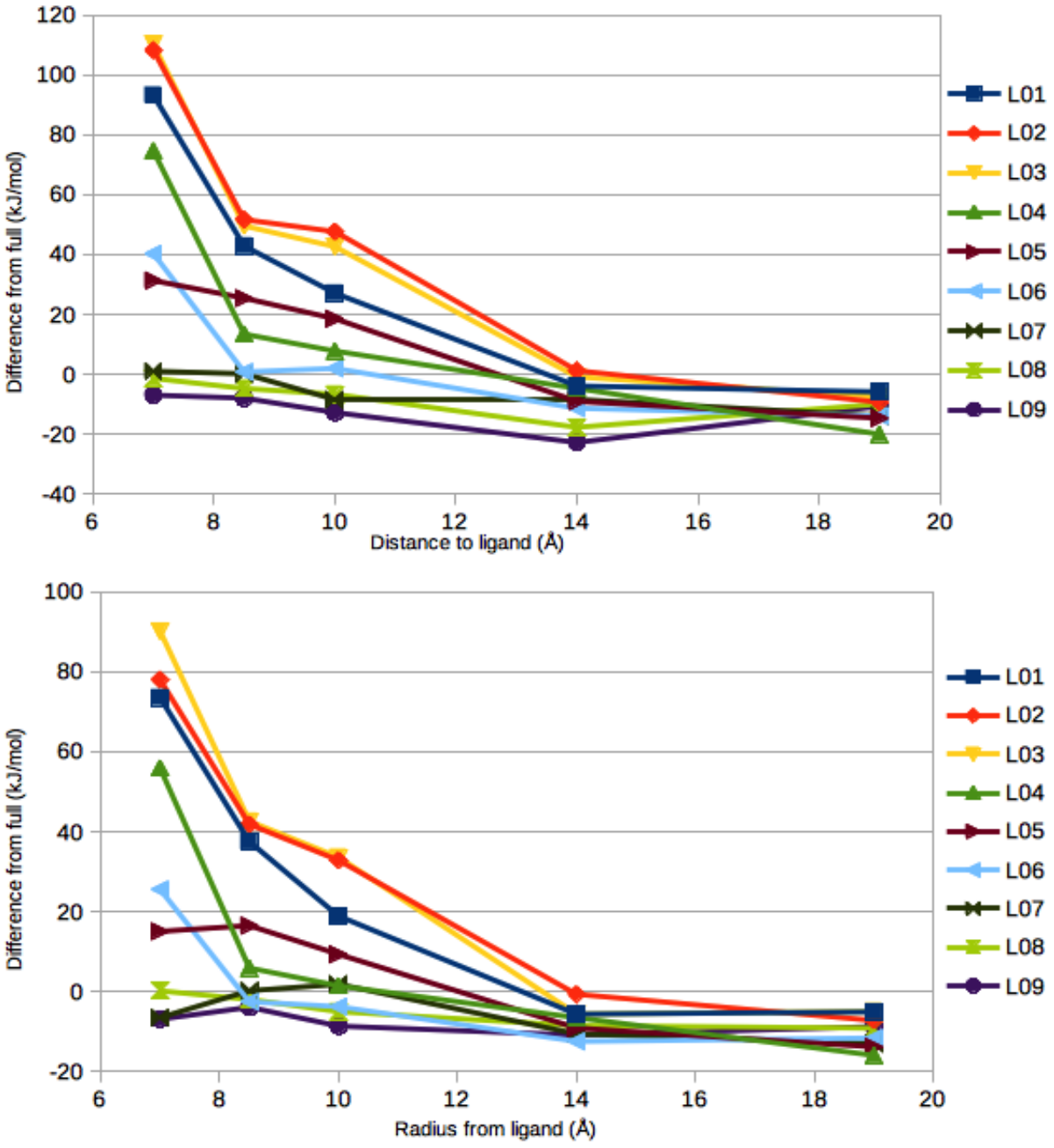

Difference from full system (kJ/mol)
10
5
0
-5
-10
6
8
10
12
14
16
18
20
Radius from ligand (Å)
L01
L02
L03
L04
L05
L06
L07
L08
L09

Figure 6. Effect of system truncations for MM/GBSA calculations based on MD simulations run on truncated systems started from the solvated crystal structure with fully solvated octahedral systems without caps (T4). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius. The top figure shows total energies, the middle figure E_{vdW}, and the lower figure E_{el}, G_{GB}, and $E_{el} + G_{GB}$ (full, dotted, and dashed lines, respectively). G_{np} is always between –1 and 1 kJ/mol and is not shown.



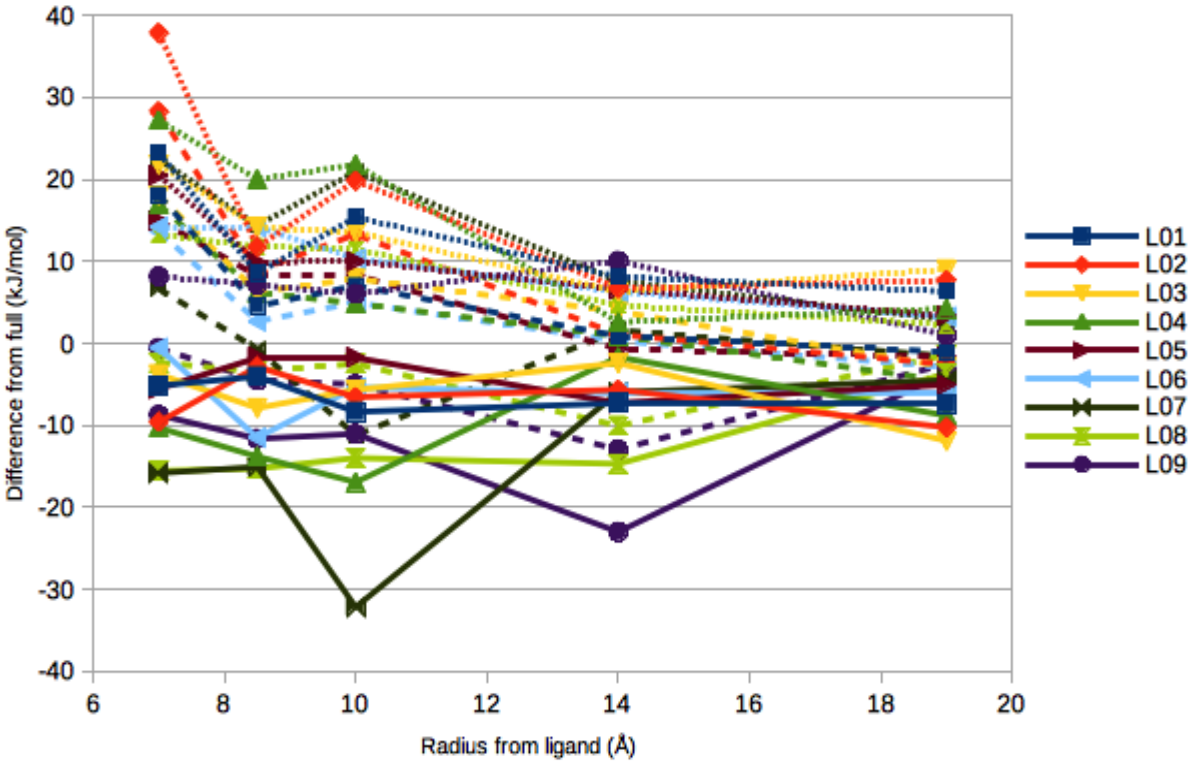

40
30
20
10
0
-10
-20
-30
-40
6
8
10
12
14
16
18
20
Difference from full (kJ/mol)
Radius from ligand (Å)
L01
L02
L03
L04
L05
L06
L07
L08
L09

Figure 7. Variation of the MM/GBSA binding affinity for L01 as a function of the rotation of the ligand around an axis through the phenol ring with the T4 truncation scheme and R = 7 Å. The two insets show typical structures from each of the two clusters.

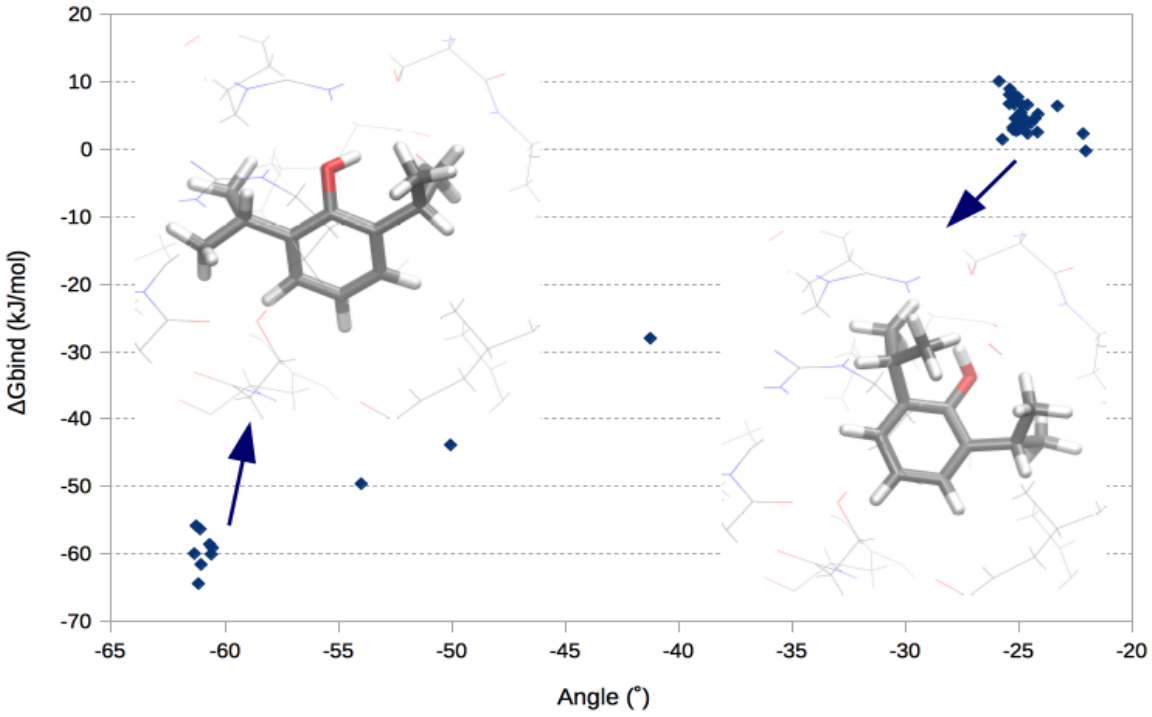


Figure 8. Effect of the number of fixed atoms in the MM/GBSA calculations based on MD simulations run on truncated systems started from the solvated crystal structure with fully solvated octahedral systems without caps (T4). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand. Results are given for R = 7 Å and residues with no atom within 3 or 2 Å of the ligand fixed. The first two curves show the total ΔG_{bind}, whereas the latter two curves show the E_{vdW} component.

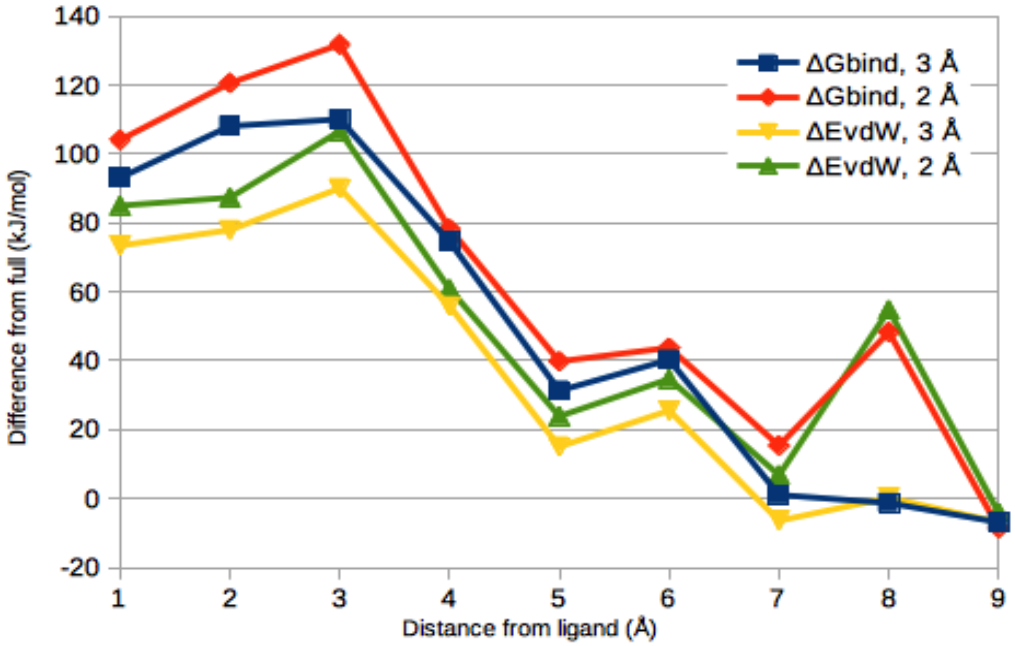
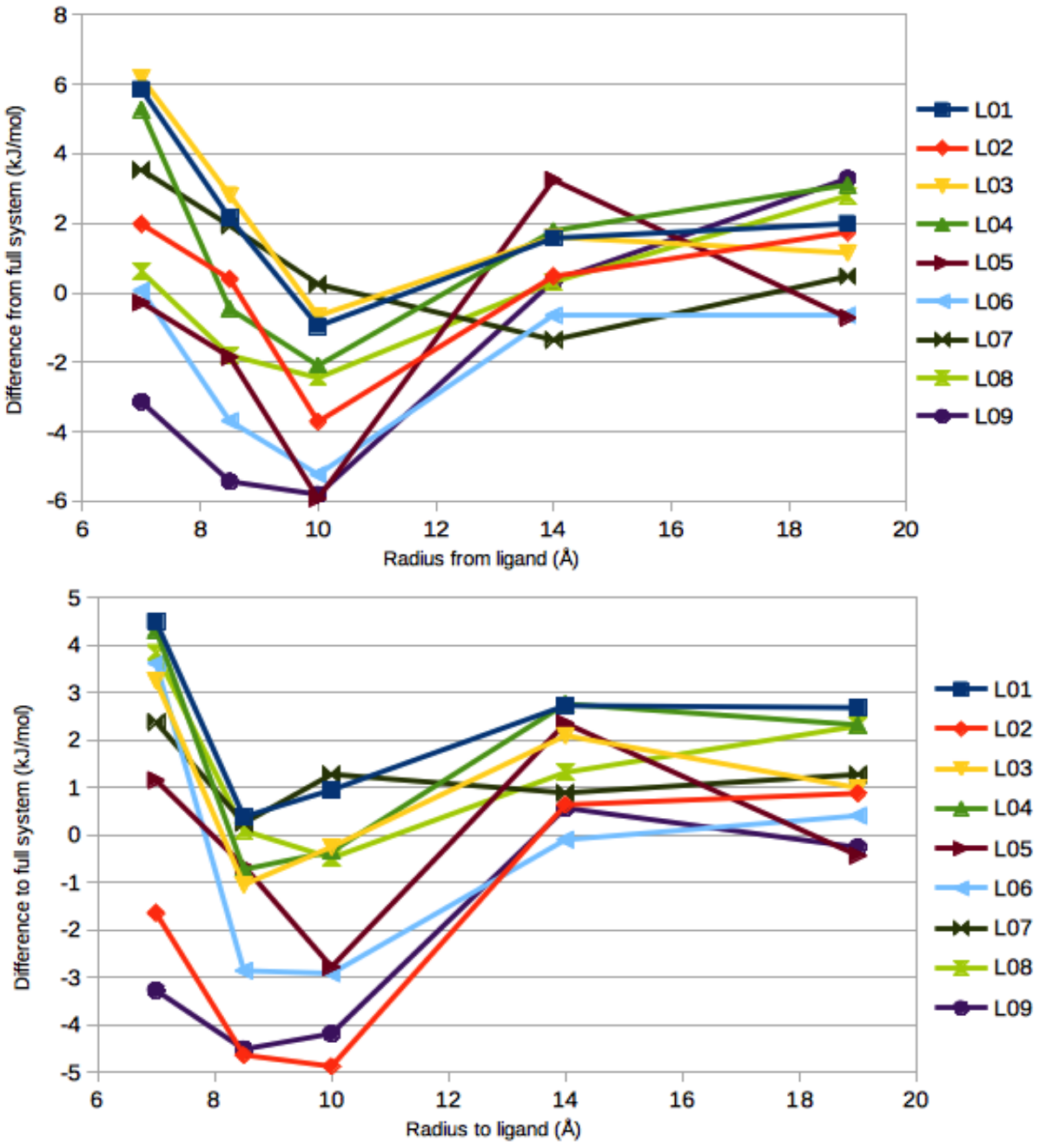


Figure 9. Effect of system truncations for MM/GBSA calculations based on MD simulations run on truncated systems started from the solvated crystal structure with fully solvated octahedral systems without any fixed atoms (T4f). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius. The top figure shows total energies, the middle figure E_{vdW}, and the lower figure E_{el}, G_{GB}, and $E_{el} + G_{GB}$ (full, dotted, and dashed lines, respectively). G_{np} is always between –1 and 1 kJ/mol and is not shown.



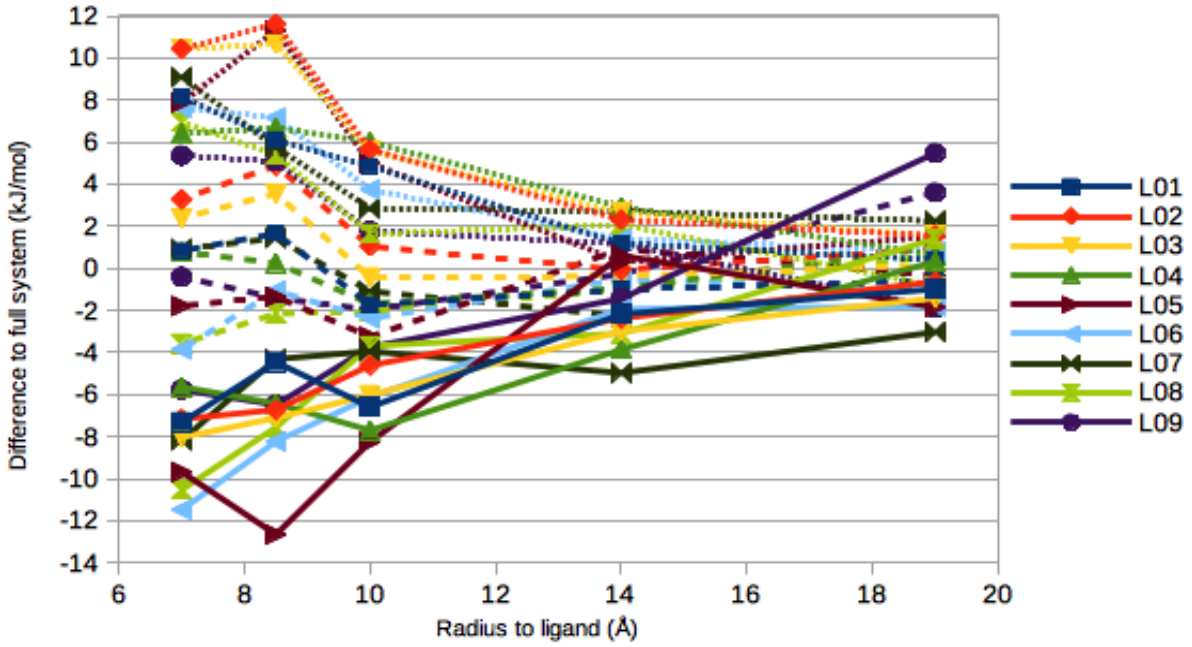

Difference to full system (kJ/mol)
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
6
8
10
12
14
16
18
20
Radius to ligand (Å)
L01
L02
L03
L04
L05
L06
L07
L08
L09

Figure 10. Effect of system truncations for MM/GBSA calculations on p38 based on MD simulations run on truncated systems started from the solvated crystal structure with fully solvated octahedral systems with (T4, a) or without any fixed atoms (T4f, b). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius.

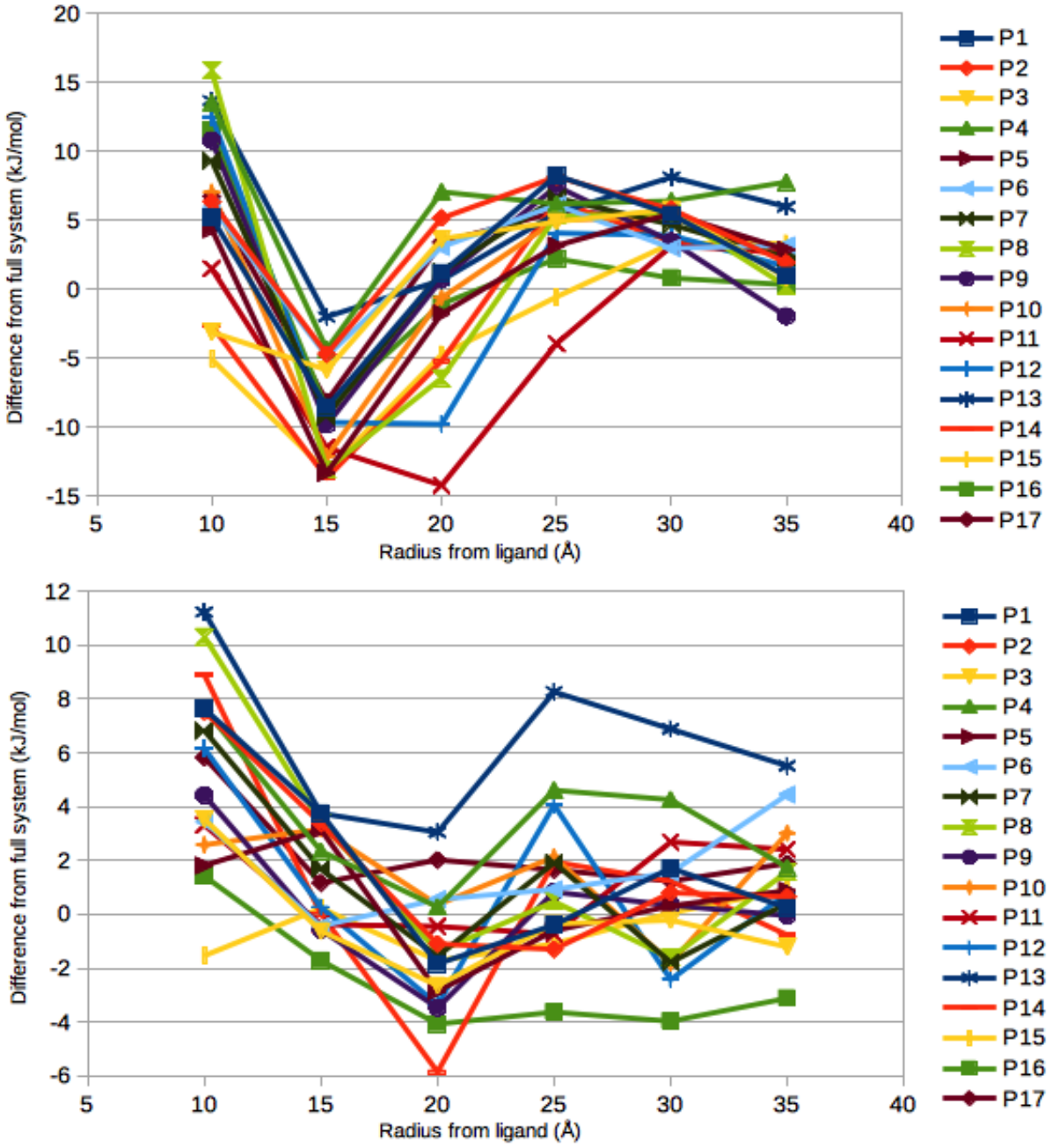
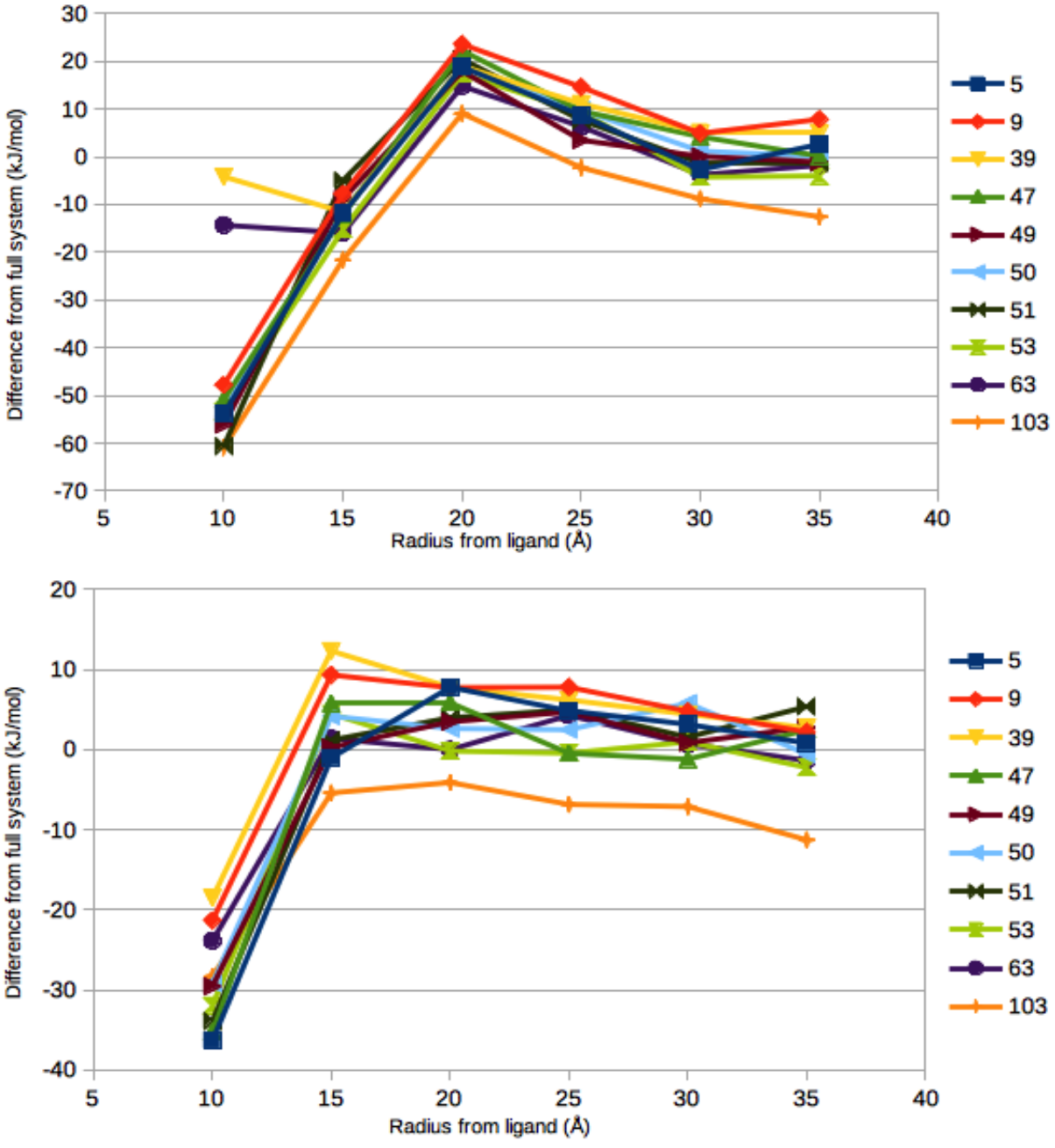


Figure 11. Effect of system truncations for MM/GBSA calculations on fXa based on MD simulations run on truncated systems started from the solvated crystal structure with fully solvated octahedral systems with (T4, a) or without any fixed atoms (T4f, b). The difference in total MM/GBSA absolute binding energy between the full and the truncated calculation is shown (kJ/mol) for each ligand as a function of the truncation radius.



TOC image and text

We have studied whether calculations of the binding free energy of small ligands to a protein by the MM/GBSA approach can be sped up by including only a restricted number of atoms close to the ligand. If the protein is truncated before the molecular dynamics simulations, quite large changes are observed for the calculated binding energies, but the results are strongly improved if no atoms are fixed in the simulations. If energies are calculated on snapshots that are truncated after the simulation, the average error is less than 1 kJ/mol at 8.5 Å truncation radius.