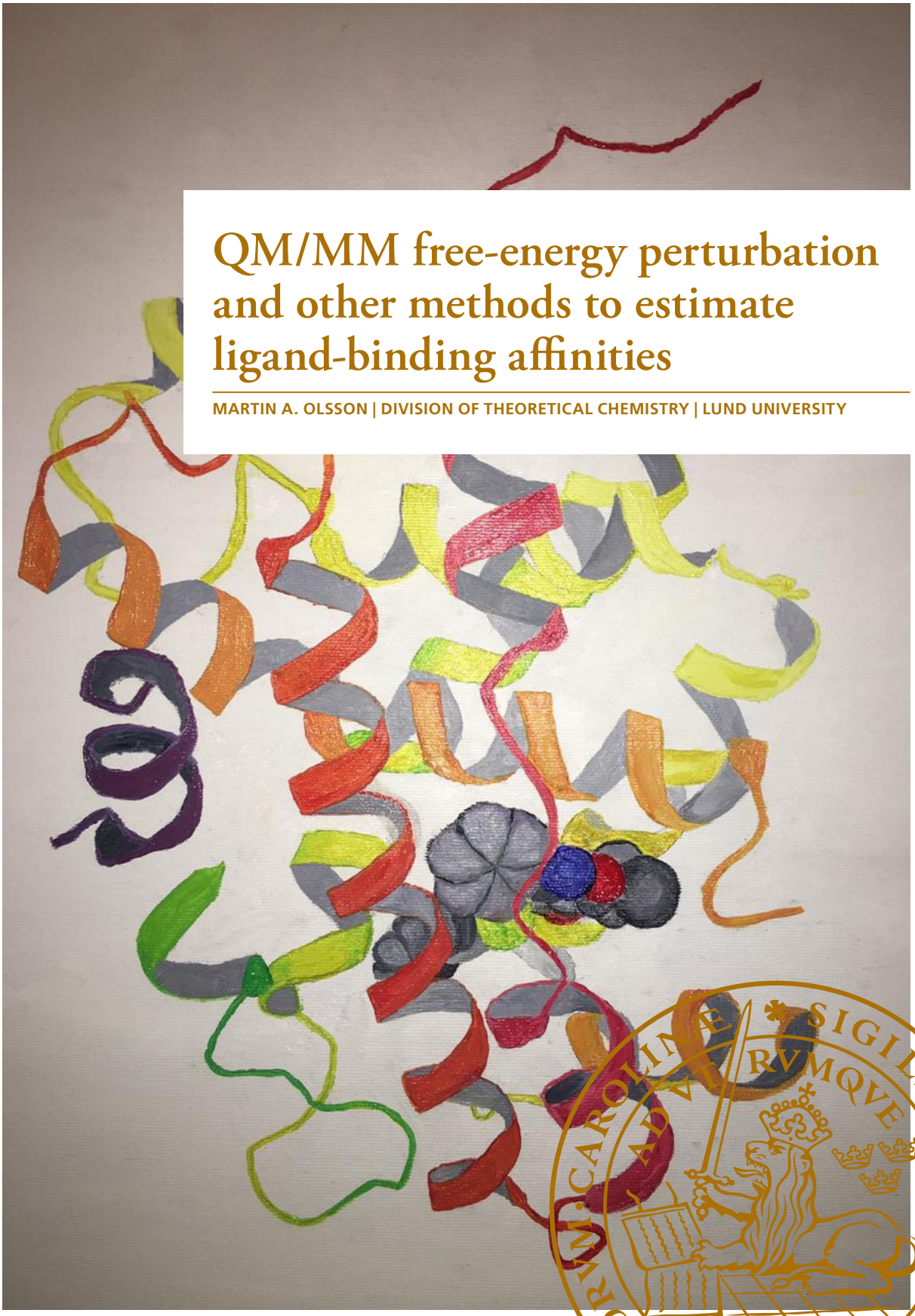
QM/MM free-energy perturbation and other methods to estimate ligand-binding affinities

Olsson, Martin

2018

Document Version:
Publisher's PDF, also known as Version of record



QM/MM free-energy perturbation and other methods to estimate ligand-binding affinities

MARTIN A. OLSSON | DIVISION OF THEORETICAL CHEMISTRY | LUND UNIVERSITY

QM/MM free-energy perturbation and other methods to estimate ligand-binding affinities

Martin A. Olsson

Division of Theoretical Chemistry
Lund University, Sweden

LUND
UNIVERSITY

DOCTORAL THESIS

Faculty opponent

Prof. Bernard R. Brooks

Grading committee

Prof. Johan Åqvist (Uppsala University)

Assoc. Prof. Frank Jensen (Aarhus University)

Prof. Göran Widmalm (Stockholm University)

By due permission of the Faculty of Science, Lund University, Sweden.

To be defended 23 March 2018 in lecture Hall A

at Chemical Centre, Lund University

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION
	Date of issue 2018-02-12
Author(s) Martin A. Olsson	Sponsoring organization Knut and Alice Wallenberg foundation

Title and subtitle
QM/MM free-energy perturbation and other methods to estimate ligand-binding affinities

Abstract

Experimental drug discovery is very time-consuming, risky and comes at a huge cost, typically several billion USD per drug. Even though decades of experimental drug discovery have provided cures of many diseases, there are still diseases for which there is no effective drug available. If drug development could be performed by theoretical and computational methods it would be of great advantage to humanity and is likely to accelerate drug discovery. One of the most promising computational methods is free-energy perturbation, which can provide estimates of protein–ligand binding affinities based on a molecular mechanics (MM) potential. Due to limitations of empirical potential-energy functions used to describe molecular interaction, there has been some interest to perform free-energy perturbation instead at the quantum-mechanics (QM) level of theory. To avoid the cost of performing sampling at the QM/MM level of theory, thermodynamic cycles can be employed. For this purpose, MM→QM free-energy perturbations in method space are required, but early applications have had convergence problems. In this thesis, different approaches to converge QM/MM free-energy perturbations in method space are developed and compared to other methods to estimate protein–ligand binding affinities. Methods to obtain QM/MM energies by performing MM→QM free-energy perturbations using thermodynamic cycles are compared to direct alchemical free-energy perturbation with a QM/MM Hamiltonian. Moreover, alternative methods to improve free-energy perturbations at the MM level of theory by charge perturbations are assessed, as well as the use of QM/MM optimised structures. Furthermore, we study also the binding entropy contribution to ligand-binding affinities for the cancer target galectin-3. QM/MM free-energy perturbation calculations in this thesis have been converged to a precision of 1 kJ/mol. The calculated free energies agree with experimental data to within 4–6 kJ/mol, which allows for a proper ranking of lead candidates. For diastereomeric inhibitors of galectin-3, both qualitative and quantitative agreement between experimental and converged binding entropy contributions to binding affinities have been obtained with a precision of ~5 kJ/mol.

Key words
Computer-aided drug design, protein–ligand binding, molecular dynamics, free-energy perturbation, QM/MM, convergence

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language English
ISSN and key title		ISBN 978-91-7422-577-8
Recipient's notes	Number of pages: 167	Price
	Security classification	

QM/MM free-energy perturbation and other methods to estimate ligand-binding affinities

Martin A. Olsson

LUND
UNIVERSITY

Front cover:
Canvas oil painting, 'Protein–ligand binding'. © Maria Serbriakova

Back cover:
Drawing 'QM/MM partitioning'. © Xiaoyi Yan and Martin A. Olsson

Funding information:
This work was financially supported by the Knut and Alice Wallenberg foundation.

Faculty of Science
Division of Theoretical Chemistry

ISBN 978-91-7422-577-8

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

MADE IN SWEDEN

For Yanzi

Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I. **Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level.**
Martin A. Olsson, Pär Söderhjelm, Ulf Ryde, *Journal of Computational Chemistry*, **2016**, 37, 1589–1600

II. **Comparison of methods to obtain ligand-binding free energies with QM/MM methods.**
Martin A. Olsson, Ulf Ryde, *Journal of Chemical Theory and Computation*, **2017**, 13, 2245–2253

III. **Relative ligand-binding free energies calculated from multiple short QM/MM MD Simulations.**
Casper Steinmann, Martin A. Olsson, and Ulf Ryde, *Journal of Chemical Theory and Computation, submitted*, **2018**

IV. **Binding free energies in the SAMPL6 octa-acid host–guest challenge calculated with MM and QM methods.**
Octav Caldararu, Martin A. Olsson, Majda Misini Ignjatović, Meiting Wang, Ulf Ryde, Manuscript, **2018**

V. **Binding affinities of the farnesoid X receptor in the D3R Grand Challenge 2 estimated by free-energy perturbation and docking.**
Martin A. Olsson, Alfonso. T. García-Sosa, Ulf Ryde, *Journal of Computer-Aided Molecular Design*, **2018**, 32, 221–224

VI. **Detailed characterization of the binding of diastereomeric ligands to galectin-3.**
Maria Luisa Vertaramo, Olof Stenström, Majda Misini Ignjatović, Octav Caldararu, Martin A. Olsson, Francesco Manzoni, Hakon Leffler, Esko Oksanen, Derek T. Logan, Ulf J. Nilsson, Ulf Ryde, and Mikael Akke, Manuscript, **2018**

List of papers not included in this thesis

Binding free energies in the SAMPL5 octa-acid host–guest challenge calculated with DFT-D3 and CCSD(T).
Octav Caldararu, Martin A. Olsson, Christoph Riplinger, Frank Neese, Ulf Ryde, *Journal of Computer-Aided Molecular Design*, **2016**, 31, 87–106

Large-scale test of the influence of charges on ligand-binding affinities calculated by free-energy simulations.
Majda M. Ignjatović, Martin A. Olsson, Ulf Ryde, Manuscript, **2018**

Structure and thermodynamics of ligand–fluoride interactions with galectin-3 backbone- and side-chain amides.
Rohit Kumar, Kristoffer Peterson, Martin A. Olsson, Johan Wallerstein, Hakon Leffler, Mikael Akke, Ulf Ryde, Ulf Nilsson, Derek T. Logan, Manuscript, **2018**

List of contributions to the papers

I. I performed all the free-energy simulations and post-processing for obtaining QM/MM energies. I tested the cumulant expansion ssEAc for the exponential averaging. I wrote the first draft of the manuscript.

II. I designed the coupling parameter expansion in method space for the reference-potential method with the AMBER software. I performed free-energy simulations in method space for all perturbations. I wrote the first draft of the manuscript.

III. I performed free-energy simulations for 8 of 9 host–guest complexes. I participated in writing of the manuscript.

IV. I performed all molecular dynamics simulations for the SAMPL6 ligands and derivation of missing force field parameters. I performed all free-energy simulations at the MM level and convergence analysis. I participated in writing of the manuscript.

V. I designed the major part of the free-energy perturbation mapping, convergence cycles and performed the free-energy simulations, as well as performed the charge correction calculations. I wrote the first draft of the manuscript.

VI. I performed the molecular dynamics simulations and conformational entropy calculations. I wrote the corresponding part of the manuscript.

Populärvetenskaplig sammanfattning

Proteiner är en av människokroppens byggstenar och har en mängd viktiga och nödvändiga biologiska funktioner genom att delta i kemiska reaktioner och processer. De flesta proteiner har en bindningsficka som styr proteinets biologiska funktion genom bindning av andra molekyler. Denna biologiska funktion kan i vissa fall kopplas till ett sjukdomstillstånd. Organiska molekyler som binder starkt till bindningsfickan hos ett protein kan hämma eller lindra ett sjukdomstillstånd. I sådana fall utgör den organiska molekylen ett läkemedel. En kritisk faktor som avgör om en organisk molekyl är bättre än en annan organisk molekyl på att binda till ett protein är bindningsaffiniteten mellan proteinet och den organiska molekylen.

Hittills har den mesta läkemedelsforskningen bedrivits genom experimentella studier med syntes av nya organiska molekyler för mätning av deras bindningsaffiniteter till ett protein. Experimentell läkemedelsforskning är mycket tidskrävande, riskfylld och kostsam, vanligtvis tiotals miljarder kronor per läkemedel. Även om årtionden av experimentell läkemedelsupptäckt har tillhandahållit botemedel för olika sjukdomar, finns det fortfarande sjukdomar för vilka det inte finns något effektivt läkemedel. Det skulle vara till stor fördel för mänskligheten om sådan läkemedelsutveckling kunde utföras med teoretiska och datoriserade metoder, vilket sannolikt skulle påskynda upptäckten av nya läkemedel. I två decennier har den förbättrade prestandan för datorer och framsteg i att ta fram 3D-modeller av proteiner möjliggjort utveckling av läkemedel med hjälp av teoretiska metoder.

Denna avhandlings första del handlar om hur man med hjälp av kvantmekanik kan förbättra metoder att beräkna fria energier, något som konventionellt görs genom att simulera proteiner och läkemedelsmolekyler med klassisk mekanik och empiriska potentialfunktioner. Denna utveckling är delvis baserad på metoder att kombinera kvantmekanik och molekylmekanik, som ursprungligen utvecklades av nobelpristagaren Arieh Warshel.

Avhandlingens andra del handlar om att förbättra konventionella metoder att beräkna fria energier genom att modellera biomolekyler med molekylmekanik-baserade kraftfält. Vi har bl.a. utvecklat korrektioner för att beräkna skillnaden i bindningsstyrka för två läkemedelsmolekyler som skiljer sig i nettoladdning, för vilket det hittills inte har funnits några lättillgängliga metoder.

Avhandlingens tredje del handlar om hur man kan tillämpa teoretiska metoder för att förklara hur små strukturella skillnader mellan läkemedelsmolekyler bidrar till bindningsaffiniteter genom beräkningar av konformationell entropi för cancerläkemedel.

Popular science summary

Proteins are one of the building blocks of the human body. They have a variety of important and necessary biological functions, participating in chemical reactions and processes. Most proteins have one or several binding pockets and binding of other molecules controls the biological function of the protein. This function may in some cases be linked to a disease. With organic molecules that effectively bind to the binding pocket of a protein, the biological function of the protein may be modified. In such cases, the organic molecule may become a drug. A critical factor determining whether an organic molecule is better than another organic molecule to bind to a protein is the difference in binding affinities between the protein and the organic molecules.

So far, most drug discovery has been conducted through experimental studies, involving synthesis of new organic molecules and measurement of their binding affinities to a protein. Experimental drug research is very time-consuming, risky and comes at an immense cost, usually several billion dollars per drug. Although decades of experimental drug discovery have provided cures for many diseases, there are still diseases for which no effective drug is available. It would be of great benefit to humanity if drug discovery could be performed with theoretical and computational methods and it would likely accelerate the discovery of new drugs. The improved performance of computers and advances in 3D-modelling of proteins during the latest two decades have enabled the development of drugs using theoretical methods.

The first part of this dissertation deals with how quantum mechanics can be used to improve calculated binding free energies, conventionally obtained by simulating proteins and drug molecules according to classical mechanics and with empirical potential-energy functions. This is partly based on methods to combine quantum mechanics and molecular mechanics, originally developed by the Nobel laureate Arieh Warshel.

The second part of the dissertation describes how to improve conventional methods to calculate free energies by modelling biomolecules with molecular mechanics-based force fields. We implement corrections to calculate the difference in binding affinity of two drug molecules that differ in net charge, for which there have been no readily available methods so far.

The third part of the dissertation shows how theoretical methods can be used to explain how small structural differences of drug molecules contribute to binding affinities through calculations of conformational entropies for two cancer drugs.

1 Introduction

1.1 Computational drug discovery

Since the 1980s, the trend in increasing computer power together with advances in protein crystallography and drug discovery have suggested that drugs one day may be developed by computational methods. Consequently, many methods have been developed for this aim[1], e.g. docking[2], MM/PBSA (molecular mechanics combined with Poisson–Boltzmann and solvent-accessible surface-area solvation)[3,4] and linear interaction energy methods[5]. However, the most accurate results are typically obtained with free-energy perturbation (FEP) and other free-energy simulation methods[6,7]. These are based on strict statistical mechanics theory and should in principle be limited only by the accuracy of the energy function employed and the sampling of the phase space.

1.2 Protein–ligand binding

In this thesis, I have studied the binding of a protein (P) and a ligand (L) to form a protein–ligand complex PL. The process is described by the chemical reaction:
P + L $\rightarrow$ PL. It is characterized by the binding constant K_{bind}, which is defined as:

$$K_{\text{bind}} = \frac{[\text{PL}]}{[\text{P}][\text{L}]} \qquad (1.1)$$

The binding free energy can then be obtained from the following equation:

$$\Delta G_{\text{bind}} = -RT \ln K_{\text{bind}} C_{\text{ref}} \qquad (1.2)$$

where K_{bind} is the binding constant, C_{ref} is the reference concentration (1 M), R is the gas constant and T is the temperature. The binding free energy is the measure for the driving force of the binding reaction.

Thus, the relative free-energy of two ligands 1 and 2 is readily obtained by the following expression:

$$\Delta G_{\text{bind}}^{1\to 2} = RT \ln \frac{K_{\text{bind}}^{(1)}}{K_{\text{bind}}^{(2)}} \qquad (1.3)$$

The binding free energy of a protein–ligand complex can be divided into two components, the enthalpy and entropy:

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \qquad (1.4)$$

The enthalpy, ΔH, describes the energetics of the interaction between the ligand and the protein or water, whereas the entropy, ΔS, is related to the probability to observe certain conformations, which is related to the conformational flexibility and variability of the complex.

The entropy of a protein–ligand complex can be written as:

$$S = k_{\text{B}} \ln \Omega \qquad (1.5)$$

where k_{B} is Boltzmann's constant and Ω is the number of conformational states of the complex.

1.3 Accuracy and precision

In this thesis, the accuracy of calculated binding estimates is measured by their deviation from the experimental values. The precision of a free-energy calculation is the statistical reproducibility of the estimate, $\Delta G_{\text{i}}^{\text{calc}}$, which is described by the mean, μ, and the standard deviation of the sampling distribution corrected for the square

root of the sample size, i.e. the standard error, σ, of the mean. Thus, the accuracy is the deviation of μ from the experimental value, $\Delta G_{\mathrm{i}}^{\mathrm{exp}}$, and the precision is the reproducibility of μ, described by σ.

When comparing theoretical predictions of binding free energies, the mean average deviation (MAD) to experiment is a useful measure of the accuracy and it is defined as:

$$\mathrm{MAD} = \frac{1}{N}\sum_{i}^{N}\left|\Delta G_{\mathrm{i}}^{\mathrm{calc}} - \Delta G_{\mathrm{i}}^{\mathrm{exp}}\right| \qquad (1.6)$$

where N is the number of samples. Another quality metric is the correlation between the calculated and experimental affinities, measured by Pearson's correlation coefficient, r^2, defined as:

$$r^2 = \left(\frac{\mathrm{cov}(\Delta G_{\mathrm{i}}^{\mathrm{calc}}, \Delta G_{\mathrm{i}}^{\mathrm{exp}})}{\sigma_{\Delta G^{\mathrm{calc}}}\,\sigma_{\Delta G^{\mathrm{exp}}}}\right)^2 \qquad (1.7)$$

where $\mathrm{cov}(\Delta G_{\mathrm{i}}^{\mathrm{calc}}, \Delta G_{\mathrm{i}}^{\mathrm{exp}})$ is the covariance between $\Delta G_{\mathrm{i}}^{\mathrm{calc}}$ and $\Delta G_{\mathrm{i}}^{\mathrm{exp}}$, and $\sigma_{\Delta G^{\mathrm{calc}}}$ and $\sigma_{\Delta G^{\mathrm{exp}}}$ are the corresponding standard deviations.

A third metric is the Kendall's τ, defined as:

$$\tau = \frac{n_{\mathrm{c}} - n_{\mathrm{d}}}{n_0} \qquad (1.8)$$

where n_{c} is the number of concordant pairs (i.e. pairs of ligands that are predicted with the correct order of affinities), n_{d} is the number of discordant pairs (incorrect predicted order) and n_0 is the number of pairs of ligands for which the experimental

affinities are not identical. Both r and Kendall's τ range from –1 to 1. In the context of free-energy calculations, a Kendall's τ close to 1 shows that the ranking of the binding free energies is correct.

A change in one order of magnitude in the binding constant (e.g. from K_{bind} = 1 nM to 10 nM) is equivalent to a room temperature change of 5.7 kJ/mol (i.e. RT ln 10) in the binding free energy. Hence, a useful theoretical method to estimate rigorous ligand-binding free energies in principle requires a MAD < 5.7 kJ/mol. Therefore, the precision of free-energy estimates should be ~1 kJ/mol to give statistically meaningful results.

The aim with this thesis has been to improve the FEP method by quantum-mechanics calculations (Section 2.1) and in particular to assess the precision (with methods described in Section 5). Section 3 discusses the sampling method of molecular dynamics simulations that the FEP calculations (Section 4) in this thesis are based on. Other methods that have been used to obtain estimates of ligand-binding affinities are entropy calculations (Section 3.2) and other QM/MM methods to estimate ligand-binding affinities (Sections 2.3 and 6).

2 Energy functions

2.1 Ab initio quantum-chemical methods

In the *ab initio* methods (ab initio meaning 'from the beginning' i.e. that no empirical data are needed), the Schrödinger equation

$$\hat{H}\Psi = E\Psi \qquad (2.1)$$

is solved. Here, $\hat{H}$ is the (Hamilton) operator for the system of interest, i.e. the kinetic energy of all particles involved plus the potential energy, E is the energy (a scalar) and Ψ is the solution, the wavefunction from which all the properties of the system can be derived. Ψ is a function of the three Cartesian coordinates of all particles in the system. The equation is simple to formulate, but impossible to solve exactly, except for a few very simple systems, e.g. the hydrogen atom and the H_2^+ ion.

Therefore, approximate methods have to be used. Contemporary quantum chemistry involves the solution of this equation using various approximations. These calculations can be performed at many levels of theory. The simplest ab initio quantum mechanical (QM) approach is the Hartree–Fock method (HF). In this, the electrons are assumed to move in the average field of all the other electrons. This is a rather crude approximation, which can be improved, e.g. by perturbation theory or series expansion. Most QM methods solve the Schrödinger equation by expanding the wavefunction in a set of known functions, a basis set. In principle, the larger this basis set is, the more accurate will the results be.

An alternative to the ab initio methods is density-functional theory (DFT). In 1964, Hohenberg and Kohn proved that the energy of a molecule is determined by its electron density[8]. The latter is a function of only the three Cartesian coordinates and therefore much simpler than the wavefunction. Still, there is a one-to-one correspondence between the wavefunction and the electron density. Unfortunately, the functional relation between the electron density and the energy is not known. Therefore, there exist a large number of variants of DFT. However, the best DFT methods give results that are much better than HF at a similar cost and therefore DFT is currently the most popular QM method. DFT calculations were employed in Paper IV.

The prime problem with QM methods is their very large computational cost. Currently, a reasonably accurate DFT energy can be calculated for almost 1000 atoms within 24 hours in a single-core computer, but DFT can hardly be used for a full middle-sized protein or for extensive sampling. QM methods can be sped up by a factor of ~1000 by replacing most of integrals involved in the solution of the Schrödinger equation by empirical parameters. This gives rise to the semiempirical QM methods. Again, there are many such methods at many levels of approximation. They are often supplemented by empirical expressions for dispersion, hydrogen bonds and halogen bonds. Semiempirical QM calculations were employed in Papers I–IV.

2.2 Molecular Mechanics

Semiempirical methods are still too costly to allow extensive sampling for a full solvated protein. Therefore, more approximate methods are needed to treat such molecules. Molecular-mechanics (MM) methods ignore the electronic structure of matter and describe molecules as balls connected by springs. Such an approximation

is far from the true picture of a molecule, but practical in terms of the low computational cost.

In molecular mechanics, the potential energy of a molecule is given by an empirical energy function (i.e. a mathematical expression that gives the energy as a function of the Cartesian coordinates of all atoms), which for a protein typically takes the following form:

$$\begin{aligned} U = & \sum_i \sum_{j \neq i} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \\ & + \sum_i \sum_{j \neq i} \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}} \\ & + \sum_{\text{bonds}} k_i^b (r_i - r_0)^2 \\ & + \sum_{\text{angles}} k_i^a (\theta_i - \theta_0)^2 \\ & + \sum_{\text{dihedrals}} k_i^\phi \left[1 + \cos(n_i \phi_i + \delta_i) \right] \end{aligned} \qquad (2.2)$$

which comprises terms describing the internal energy of a molecule (bonds, angles and torsions; last three terms) and its interaction energy that arises from van der Waals and Coulomb forces. The first term,

$$\sum_i \sum_{j \neq i} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \qquad (2.3)$$

is the Lennard–Jones potential between all pairs of atoms, *i* and *j*, in the molecule. It consists of two terms: a repulsion between atoms, caused by their electron cloud penetration, i.e. steric hindrance, and an attractive term based on the London dispersion; $-\varepsilon_{ij}$ is the depth of the potential energy curve and a measure of how

strongly two atoms attract each other, σ_{ij} is the distance at which the inter-particle potential is zero, which describes how close the two atoms can get, and r_{ij} is the distance between the two atoms. The second term,

$$\sum_i \sum_{j \neq i} \frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}} \qquad (2.4)$$

is the Coulomb electrostatic interaction energy, which describes the attraction or repulsion of two atoms based on their partial atomic charges q_i and q_j. The third term,

$$\sum_{\text{bonds}} k_i^b (r_i - r_0)^2 \qquad (2.5)$$

describes the harmonic bond energy as a spring potential, where k_i^b is the spring constant for the bond type b, r_i is the distance between the two atoms and r_0 is the equilibrium bond length.

Similarly, the term

$$\sum_{\text{angles}} k_i^a (\theta_i - \theta_0)^2 \qquad (2.6)$$

describes a harmonic potential for the bond angles, where k_i^a is the spring constant for the angle a, θ is the angle between the three atoms and θ_0 is the equilibrium angle.

The last term is the potential for dihedral angles,

$$\sum_{\text{dihedrals}} k_i^\phi \left[1 + \cos(n_i \phi_i + \delta_i)\right] \qquad (2.7)$$

where k_i^ϕ is the spring constant for the dihedral angle ϕ_i, n_i is the periodicity of the dihedral and δ_i is the phase shift. Classical potential-energy functions like the one in Eqn. 2.2 are pairwise additive. They contain a very large number of parameters (all

ε_{ij}, σ_{ij}, q_i, k_i^b, r_0, k_i^a, θ_0, k_i^ϕ, n_i, and δ_i). Each of them has to be determined from empirical or computational (QM) data. There exists a large number of different force fields of different sophistication. We have employed the AMBER ff14SB force field for proteins, the generalized AMBER force-field (GAFF) for organic molecules (like drug candidates) and the TIP3P or TIP4P-Ew models for water.

2.3 QM/MM

QM methods gives an accurate description of the electronic structure of atoms and molecules and are especially important in the study of chemical reactions. However, calculating the energy of every single electron of a protein appears unnecessary for the study of protein–ligand binding. Quite often, only a certain part of the protein is of central interest (e.g. the active site or the ligand-binding site) as illustrated in Figure 2.1. Then, it may be better to use an accurate QM Hamiltonian for this part and a less accurate, but faster MM Hamiltonian for the remainder of the protein, as well as a substantial amount of explicit water molecules. This gives rise to the combined QM/MM approach. In this, the total energy is obtained as:

$$E_{\text{total}} = E^{\text{QM}} + E^{\text{MM}} + E^{\text{QM/MM}} \qquad (2.8)$$

where E^{QM} is the QM energy of the QM system, E^{MM} is the MM energy of the MM systems, and $E^{\text{QM/MM}}$ is an interaction energy between the MM and QM system. QM/MM was invented by Karplus, Warshel and Levitt[9] and was the basis for their Nobel prize in chemistry 2013 “for the development of multiscale models for complex chemical systems"[10]. QM/MM potentials are used in Papers I–IV.

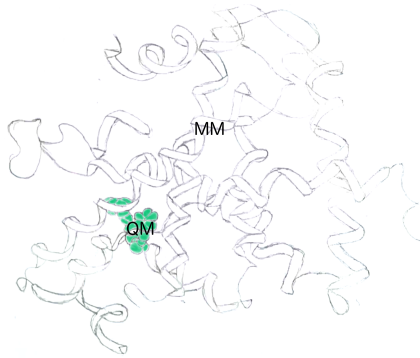


Figure 2.1. QM and MM partitioning in QM/MM depicted for protein–ligand binding with a drug molecule in the QM region.

3 Sampling

With the energy functions described in the previous section, the energy of essentially any molecule can be calculated. However, biomacromolecules (and also ligands) are flexible molecules that may attain many conformations with slightly different energies. Moreover, as can be seen in Eqn. 1.4, the binding free energy depends also on the entropy of the binding process, which is a measure of the probability of different conformations for both the protein–ligand complex and the free protein or ligand. Thus, it is of great interest to sample different conformations of the molecules involved in the binding. This can be done with different methods, e.g. by systematic search of the conformational space, Monte Carlo methods or by molecular dynamics simulations. For biomolecular systems, the latter approach is by far the most common.

3.1 Molecular Dynamics

In molecular dynamics (MD) simulations, atoms are moved according to forces obtained from the potential energy function (the forces are the first derivative of the energy with respect to the coordinates). According to Newton's second law:

$$F = ma = m\frac{d^2x(t)}{dt^2} \qquad (3.1)$$

the acceleration of an atom with mass m, can be directly obtained from the forces, F. On the other hand, the acceleration is the second time derivative of the position, $x(t)$. Thus, Eqn. 3.1 is a second-order differential equation. Such an equation can be

numerically solved iteratively by taking a small time step (0.5–2 femtoseconds), updating the velocity (which is the first time derivative of the position) in the direction of the acceleration and then moving the particles a small step in the direction of the velocities. At the new positions of all particles, the energy and forces are recalculated. By repeating this procedure numerous times, we get the time-course or trajectory of the system from which dynamic properties can be extracted and thermodynamic averages at physiological temperatures can be calculated.

Unfortunately, the method gives a very large amount of data and it is time consuming (the typical time of MD simulations is 1 ns – 10 μs and the longest simulation ever done was 1 ms[11]). In order to decrease the computational effort, bond lengths involving hydrogen atoms are often constrained to their equilibrium values. This removes the (usually uninteresting and strictly quantum-mechanical) bond vibrations without affecting other results. Typically, the so-called SHAKE[12] procedure is used. Another common way to speed up the calculations is to ignore non-bonded interactions (which are by far the most numerous interactions) for atoms more distant than a certain cut-off distance (typically at least 8 Å). This works well for van der Waals interactions, which are short-ranged, but not for electrostatic interactions, which are very long-ranged. A way to solve this problem is to use Ewald summation. It works only for periodic systems, but periodic boundary conditions (i.e. copies of the simulated box are imagined on all sides so that an atom that leaves the box at one side, simultaneously enters the box on the opposite side) are normally used to simulate infinite systems.

Atomistic molecular dynamics simulations at the MM level can be used to study fully solvated proteins of ~100 000 atoms on ps to μs time scales, which is appropriate for the biochemical processes in this thesis. Coarse-graining (i.e. replacing groups of atoms with a single interaction centre) allows molecular dynamics simulation to be used for even larger biochemical structures e.g. the entire HIV virus capsid[13] of 64

million atoms[14]. Furthermore, continuum solvation models, such as the generalised Born model[15] can drastically improve the accessible time scale of the simulations[16].

3.2 Entropy

As was shown in Eqn. 1.4, the binding free energy of a ligand depends on both the enthalpy and the entropy. Quite often compensation between the enthalpy and the entropy is observed; an increase in the enthalpy, i.e. a tighter binding of the ligand, typically gives a lower mobility of the interacting molecules and therefore leads to a decreased entropy. Naturally, this enthalpy–entropy compensation makes the optimisation of binding affinities harder.

The entropy is often further decomposed into contributions from the various parts of the complex:

$$\Delta S_{\text{bind}} = \Delta S^{\text{P}}_{\text{conf}} + \Delta S^{\text{L}}_{\text{conf}} + \Delta S^{\text{P}}_{\text{r-t}} + \Delta S^{\text{L}}_{\text{r-t}} + \Delta S_{\text{solvent}} \qquad (3.2)$$

where the entropy is subdivided into terms for the conformational entropy of the protein (P) and the ligand (L), the rotational and translational entropy for the protein and the ligand, and the solvation entropy.

There are several ways to estimate the conformational entropy from molecular dynamics simulations, e.g. normal-mode calculations and quasi-harmonic analysis[17,18]. Based on previous research in the group[19,20], we have selected to use histogramming of the dihedral distribution[21]. The protein Cartesian coordinates are converted to internal (bond, angle, and torsion) coordinates. Only the dihedral angles are necessary to calculate the conformational entropy[19].

The conformational entropy can then be written as:

$$S_{\mathrm{i}} = \frac{R}{2} - R\ln N - R\sum_{j=1}^{N} p_{\mathrm{i}}(j)\ln p_{\mathrm{i}}(j) \qquad (3.3)$$

where R is the gas constant and $p_{\mathrm{i}}(j)$ is the probability that the dihedral angle i is found in bin j, and N is the number of bins per dihedral. The first two terms of the entropy expression are normalisation factors, giving the entropy of a free rotor (R/2) for a uniform distribution (which cancel for relative entropies)[20]. Conformational entropy calculations for protein–ligand complexes were performed in Paper VI.

Experimentally, conformational entropies can be estimated from NMR measurements of structure factors. However, these experiments can currently be done only for a few chemical groups. Therefore, entropies from MD simulations are typically used to supplement the measurements. It is then of interest to check that the MD simulations are accurate. One way to calibrate the simulations is to use structure factors, which are available both from the experiments and the calculations. In paper VI, order parameters were estimated from MD simulations of the protein galectin-3 with isotropic reorientational eigenmode dynamic analysis[22].

4 Free-energy perturbation

It is the free energy that determines which chemical processes are possible and it is therefore a central property in chemistry. A large number of computational methods have been developed to estimate free energies, e.g. docking and scoring[1], MM/PBSA[3,4] and LIE[23]. However, all these methods are only approximate. The free-energy perturbation (FEP) approaches[24–28] are based on a strict statistical-mechanics theory and therefore should give the correct results, provided that the energy function is perfect and the sampling is exhaustive. In FEP, the relative free-energy difference of two systems is estimated by simulating one of the systems and calculating the energy difference between the two systems at regular intervals. The free-energy difference can be estimated by many different approaches, e.g. exponential averaging (EA), thermodynamic integration (TI)[24], Bennett acceptance ratio (BAR)[25,26] or multistate BAR (MBAR)[27]. The exponential average or Zwanzig expression was developed in 1954[28] and is still used in the field to this day. The TI approach was developed in 1935 by Kirkwood[24].

4.1 Perturbation formalism

In free-energy perturbation, the free-energy difference between two states is sought, e.g. the difference in binding affinities of two similar ligands, differing only in one functional group. According to statistical mechanics, the Helmholtz free-energy A is given by

$$A = -k_{\mathrm{B}} T \ln Q(N,V,T) \qquad (4.1)$$

where N is the number of atoms, V is the volume, T is the absolute temperature and Q is the partition function, which (in the canonical ensemble) is defined as:

$$Q = \frac{1}{N!h^{3N}} \int_{\Gamma} e^{-\beta U(x,p_x)} dx\, dp_x \qquad (4.2)$$

where h is Planck's constant, Γ is the total phase space, x is the coordinates, p_x the momentum, and $\beta = \frac{1}{k_B T}$. The momentum can be integrated out of this equation using

$$Q(N,V,T) = \frac{1}{N!\Lambda^{3N}} Z(N,V,T) \qquad (4.3)$$

where $\Lambda = \frac{h}{\sqrt{2\pi m k_B T}}$ is the thermal de Broglie wavelength and $Z(N,V,T)$ is the configurational integral defined by:

$$Z(N,V,T) = \int dx\, e^{-\beta U(x)} \qquad (4.4)$$

where U(x) is the potential energy[29].

Considering two states A and B, with the potential energy difference:

$$\Delta U = U_B - U_A \qquad (4.5)$$

then the free-energy difference is:

$$\Delta A = -\beta^{-1} \ln \frac{\int dx\, e^{-\beta U_B(x)}}{\int dx\, e^{-\beta U_A(x)}} \qquad (4.6)$$

By eliminating U_B with Eqn. 4.5, we get:

$$\Delta A = -\beta^{-1} \ln \frac{\int dx\, e^{-\beta(\Delta U(x) + U_A(x))}}{\int dx\, e^{-\beta U_A(x)}} \qquad (4.7)$$

This expression can be simplified by identifying the probability density function of the state A, that is

$$P_{\mathrm{A}}(x) = \frac{e^{-\beta U_{\mathrm{A}}(x)}}{\int dx\, e^{-\beta U_{\mathrm{A}}(x)}} \qquad (4.8)$$

which gives

$$\Delta A = -\beta^{-1} \ln\left(\int dx\; e^{-\beta \Delta U(x)}\, P_{\mathrm{A}}(x) \right) \qquad (4.9)$$

and this is usually written as

$$\Delta A = -\beta^{-1} \ln\left\langle e^{-\beta \Delta U(x)} \right\rangle_{\mathrm{A}} \qquad (4.10)$$

where $\langle ... \rangle_{\mathrm{A}}$ denotes the ensemble average over configurations sampled from the reference state A. This is the exponential average or Zwanzig equation.

A cumulant expansion of the free energy ΔA to second order gives[30]:

$$\Delta A = \left\langle \Delta U \right\rangle_{\mathrm{A}} - \frac{\beta}{2}\left(\left\langle \Delta U^2 \right\rangle_{\mathrm{A}} - \left\langle \Delta U \right\rangle^2_{\mathrm{A}} \right) \qquad (4.11)$$

The expression can also be obtained analytically (and therefore exactly) if it is assumed that ΔU follows a Gaussian distribution. The cumulant approximation was used in Paper I for QM/MM free-energy perturbation.

In general, FEP gives converged results only if the difference between the two states is small. This is often solved by dividing the perturbation into many small steps, employing a coupling parameter, λ, and using the potential-energy function:

$$U(\lambda) = (1 - \lambda) U_0 + \lambda U_1 \qquad (4.12)$$

where U_0 and U_1 are the potential functions for the two states.

Another method to calculate free-energy differences is thermodynamic integration (TI). In this approach, the average of the derivative of the potential energy with

respect to λ is cumulated during a set of simulations and the free-energy difference can then be calculated by integration:

$$\Delta A = \int_0^1 \left\langle \frac{\partial U(\lambda, x)}{\partial \lambda} \right\rangle_\lambda d\lambda \qquad (4.13)$$

An advantage with TI is that convergence problems caused by large changes in the energy are shown as jumps in the derivative $\frac{\partial U}{\partial \lambda}$, in analogy with numerical quadrature.

A third approach to calculate free energies is the Bennett acceptance ratio (BAR) method. In this, the free-energy difference between two states is estimated from:

$$e^{-\beta \Delta A} = \frac{f\left(\beta(\Delta U - C)\right)}{f\left(\beta(-\Delta U + C)\right)} + C \qquad (4.14)$$

where $f(x)$ is the Fermi function, defined by:

$$f(x) = \left(1 + e^{\beta x}\right)^{-1} \qquad (4.15)$$

and C is a constant. An iterative procedure is applied to find a value of C that makes the first term of the right-hand side of Eq. 4.14 vanish. It can be proved that BAR is the statistical estimator with the smallest variance.

Relative binding free energies $\Delta\Delta G_{\text{bind}}$ between two ligands, L_0 and L_1 can be estimated using a thermodynamic cycle (Figure 4.1) that relates $\Delta\Delta G_{\text{bind}}$ to the free energy of alchemically transforming L_0 into L_1 when they are either bound to the protein, $\Delta G^{\text{b}}_{0\rightarrow 1}$, or free in solution, $\Delta G^{\text{f}}_{0\rightarrow 1}$. Since the free energy is a state function and the path through the cycle should be zero, the free energy is obtained through the expression:

$$\Delta\Delta G_{\text{bind}}^{0\to1} = \Delta G_{\text{bind}}^{(1)} - \Delta G_{\text{bind}}^{(0)} = \Delta G_{0\to1}^{\text{b}} - \Delta G_{0\to1}^{\text{f}} \qquad (4.16)$$

where $\Delta G^{\text{b}}_{0\to1}$ and $\Delta G^{\text{f}}_{0\to1}$ are estimated by FEP. Note that the protein does not change during the calculation of $\Delta G^{\text{f}}_{0\to1}$. Therefore, it is not needed to be included in that simulation, which instead involves only the ligand in water solution (as indicated by the schematic water boxes in Figure 4.1). Thus, it is necessary to sample the protein–ligand complex in solution and the solvated ligands, but not the actual binding event in the vertical legs, which may be affected by displaced water molecules[31] or conformational changes of amino-acid residues[32] (as is discussed in Paper VI). An alternative approach is to study the actual binding event, i.e. the dissociation path of the ligand out of the binding site of the protein (e.g. the attach-pull-release method[33]).

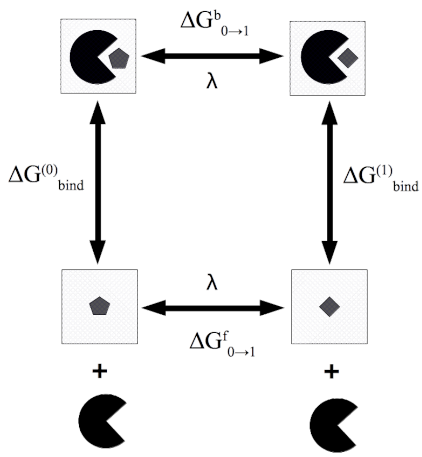


Figure 4.1. Thermodynamic cycle of protein–ligand binding with schematic depiction of a cycle formed by two protein–ligand complexes in the bound and free states. The two horizontal legs are sampled during FEP simulation (here indicated by the λ expansion). A schematic water box is depicted for each state.

To improve the convergence, relative binding free energies are estimated between a number of λ-values, according to the right-hand side of Eqn. 4.17:

$$\Delta\Delta G_{\text{bind}}^{0\to 1} = \sum_{\lambda_i=0}^{1}\left(\Delta G^{\text{b}}_{\lambda_i\to\lambda_{i+1}} - \Delta G^{\text{f}}_{\lambda_i\to\lambda_{i+1}}\right) \qquad (4.17)$$

Having calculated the relative free energy of the perturbation between ligands 1 and 2, i.e. 1→2 and the ligands 2 and 3 i.e. 2→3, one can readily obtain the free energy as:

$$\Delta\Delta G_{1\to 2\to 3} = \Delta\Delta G_{1\to 2} + \Delta\Delta G_{2\to 3} \qquad (4.18)$$

The standard error of the estimate is obtained by error propagation:

$$\sigma(\Delta\Delta G_{1\to 2\to 3}) = \sqrt{\sigma(\Delta\Delta G_{1\to 2})^2 + \sigma(\Delta\Delta G_{2\to 3})^2} \qquad (4.19)$$

4.2 Sampling errors and cycle closures

FEP estimates obtained with EA have a direction, i.e. $\Delta G_{0\to 1}$ estimated from a simulation of L_0 is not necessarily identical to $-\Delta G_{1\to 0}$ estimated from a simulation of L_1. Therefore, the difference between these two estimates, $\Delta\Delta G_{\text{EA}} = \Delta G_{0\to 1} - \Delta G_{1\to 0}$, the hysteresis, is a simple estimate of the sampling errors of the simulation (but it is independent of errors in the force field). Another way to assess the reliability of the free-energy estimates is to use convergence cycles, which connect a ligand through a series of ligands back to itself by an *in silico* simulated path. Again, the cycle-closure hysteresis gives an estimate of the statistical errors of incomplete sampling of the phase space, whereas it is independent of the errors in the force field. For a path between three ligands, 1→2→3→1, the cycle-closure free energy can be written:

$$\Delta G_{\text{closure}} = \Delta\Delta G_{1\to 2} + \Delta\Delta G_{2\to 3} + \Delta\Delta G_{3\to 1} \qquad (4.20)$$

which ideally should be zero.

4.3 Convergence criteria

FEP calculations give reliable results only if the energy distributions of the starting and end states overlap significantly, so that conformations with a high probability in one state is satisfactorily sampled also with the other state. If this is not the case, FEP will give unreliable results. This is in general not the case and this is the reason why standard FEP divides the perturbation into many small steps, using the λ coupling parameter[34]. With an increased number of λ-values, the probability-distribution overlap between neighbouring λ-values are improved (Figure 4.2), but at a higher computational cost. It is essential that the difference between the two states is not too large if the FEP calculation should give reliable results.

Likewise, the convergence of a calculated difference in binding free energy between two ligands can often be improved by employing a real or virtual ligand as an intermediate state. However, the longer free-energy path will also increase the uncertainty of the results by error propagation, so care must be taken for the statistical precision.

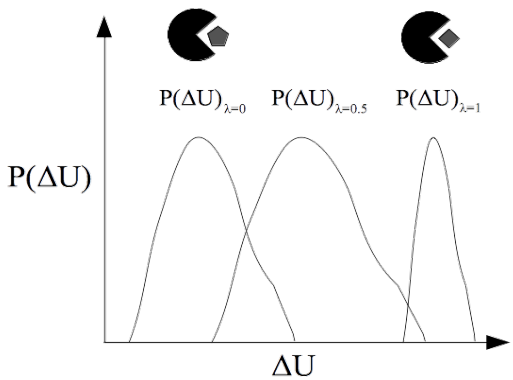


Figure 4.2. Probability distributions for the potential energy ΔU in the endpoints in FEP with one overlapping probability distribution for the inter-mediate state λ=0.5. The probability distributions for the endpoints do not overlap.

In this thesis, I have employed six measures to assess the overlap in FEP calculations:

- The Bhattacharyya coefficient[35] Ω that describes the energy distribution overlap between two λ-values.
- Wu & Kofke overlap measures of the energy probability distributions[36,37] K_{AB} .
- Wu & Kofke's bias metric[36,37] Π, which describes the important configurations of state A in state B.
- The weight of the maximum term in the exponential average[38], $w_{\max}$.
- The difference of the forward and backward exponential average estimate[34] $\Delta\Delta G_{\mathrm{EA}}$, i.e. the hysteresis of the EA calculation.
- Difference between the Bennett acceptance ratio and thermodynamic integration estimates[34,39] $\Delta\Delta G_{\mathrm{TI}}$.

4.4 Charge perturbations

Performing free-energy perturbations involving a change in the net charge of the ligands has so far been an important challenge. Such FEP give rise to considerable artefacts in the calculated energies if the simulations are performed with periodic boundary conditions and employing Ewald summations for the electrostatic interactions. In Paper V, we have implemented with the AMBER simulation package, the semi-analytic correction to $\Delta\Delta G_{\mathrm{bind}}$ developed by Rocklin et al[40,41]. The correction includes five terms:

$$\Delta\Delta G_{\mathrm{corr}} = \Delta\Delta G_{\mathrm{DSC}} + \Delta\Delta G_{\mathrm{NET}} + \Delta\Delta G_{\mathrm{USV}} + \Delta\Delta G_{\mathrm{RIP}} + \Delta\Delta G_{\mathrm{EMP}} \qquad (4.21)$$

The first term is a correction for discrete solvent effects. It is obtained from:

$$\Delta\Delta G_{\mathrm{DSC}} = -\frac{\gamma_{\mathrm{S}} Q_{\mathrm{L}}}{6\varepsilon_0} \cdot \frac{N_{\mathrm{S}}}{L^3} \qquad (4.22)$$

where N_s is the number of solvent molecules and L is the side length of the cubic simulation box, whereas Q_L is the net charge of the ligand, and ε_0 is the permittivity of vacuum, γ_s is the quadrupole moment trace of the water model, with respect to its single van der Waals interaction site. For a TIP3P water molecule, $\gamma_s = 0.0764\ e\ \mathrm{nm}^2$.

The next two terms are the corrections for periodicity-induced net-charge interactions and periodicity-induced net-charge undersolvation. They can be obtained directly from the periodicity-induced net-charge interactions as,

$$\Delta\Delta G_{\mathrm{NET}} + \Delta G_{\mathrm{USV}} = \frac{1}{\varepsilon_{\mathrm{s}}}\Delta\Delta G_{\mathrm{NET}} = -\frac{\xi_{\mathrm{LS}}}{8\pi\varepsilon_0} \cdot \frac{1}{\varepsilon_{\mathrm{s}} L}\left[(Q_{\mathrm{P}} + Q_{\mathrm{L}})^2 - Q_{\mathrm{P}}^2\right] \qquad (4.23)$$

Here, Q_P is the net charge of the protein, ε_s is the static relative dielectric permittivity of the medium ($\varepsilon_s = 97$ for the TIP3P water model) and ξ_{LS} is the cubic lattice-sum (Wigner) integration constant ($\xi_{\mathrm{LS}} \approx -2.837297$).

The fourth term, $\Delta\Delta G_{RIP}$, is a correction for residual integrated potential effects and it is calculated from:

$$\Delta\Delta G_{\mathrm{RIP}} = \frac{1}{L^3}\left[(I_{\mathrm{P}} + I_{\mathrm{L}})(Q_{\mathrm{P}} + Q_{\mathrm{L}}) - I_{\mathrm{P}} Q_{\mathrm{P}}\right] \qquad (4.24)$$

where I_P and I_L are the residual integrated potentials for the protein and ligand, respectively, in units of kJ $\mathrm{nm}^3\ \mathrm{mol}^{-1} e$.

Finally, $\Delta\Delta G_{EMP}$ is an empirical correction to the analytical result. It is given by:

$$\Delta\Delta G_{\mathrm{EMP}} = -\frac{1}{8\pi\varepsilon_0} \cdot \frac{16\pi^2}{45} \cdot \left(1 - \frac{1}{\varepsilon_{\mathrm{s}}}\right) \cdot \left[(Q_{\mathrm{P}} + Q_{\mathrm{L}})^2 - Q_{\mathrm{P}}^2\right] \cdot \frac{R^5}{L^6} \qquad (4.25)$$

This expression contains the effective solvation radius of the ligand, which is determined by the expression:

$$R_{\mathrm{L}} = \left\{ I_{\mathrm{L,\,SLV}} \Big/ \left[\frac{1}{8\pi\varepsilon_0} \cdot \frac{4\pi}{3} \cdot \left(1 - \frac{1}{\varepsilon_{\mathrm{s}}} \right) \left| Q_{\mathrm{L}} \right| \right] \right\}^{1/2} \qquad (4.26)$$

and the solvent contribution to the residual integrated potential for the ligand is calculated as:

$$I_{\mathrm{L,\,SLV}} = I_{\mathrm{L}} - I_{\mathrm{L,\,hom}} \qquad (4.27)$$

where $I_{\mathrm{L,SLV}}$ is the solvent contribution to I_{L} , where $I_{\mathrm{L,hom}}$ is the residual integrated potential obtained with the relative permittivity set to 1 everywhere. The I_{L} , $I_{\mathrm{L,hom}}$ and I_{P} values are obtained from Poisson–Boltzmann calculations of the protein or the ligand.

5 QM/MM free-energy perturbation

As mentioned above, the accuracy of FEP is essentially limited only by the sampling and the energy function. The great majority of FEP calculations are performed with a MM energy function, owing to the long MD simulations needed for proper sampling. However, recently there have been quite some interest to improve the energy function by using QM/MM calculations[42]. This can be important for many drug candidates, which often involve halogens that may be involved in halogen bonds or large π-systems that may form cation–π interactions[43], both of which are poorly described with standard MM force fields. Of course, such calculations can be performed by just replacing the MM potential with QM/MM. However, that would require long QM/MM MD simulations, which are very expensive. Therefore, such calculations have been performed only a few times and with only the ligand in the QM system, treated by a semiempirical method[44–47]. A major part of this thesis (Papers I–IV) has been devoted to developing and comparing methods to calculate FEP free energies at the QM/MM level.

5.1 MM→QM perturbation formalism

QM/MM free energies can be estimated by sampling the QM region of the system with a QM Hamiltonian and the remaining part of the system with a MM Hamiltonian. The simplest, but also most expensive, way to obtain a QM/MM free-energy estimate of an alchemical transformation is to use direct alchemical QM/MM free-energy simulation along the top path in Figure 5.1.

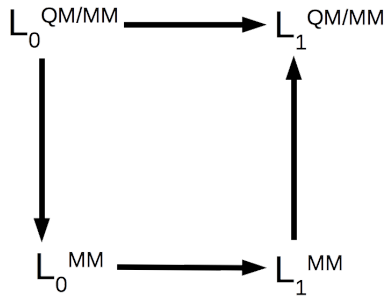


Figure 5.1. Thermodynamic cycle used in the reference-potentials methods to calculate free-energy difference at the MM and QM/MM levels of theory.

An alternative approach to avoid the sampling with the expensive QM/MM Hamiltonian is to use the reference-potential method, first suggested by Warshel[48]. Similar methods have been used by other groups[49,50] e.g. the QM/MM cycle perturbation (QTCP) approach[38] by Rod and Ryde. In these approaches, the free-energy difference is estimated at the MM level with normal FEP. At the endpoints, additional FEPs are performed in the method space from MM to QM/MM, as shown by the two vertical paths in Figure 5.1. Since the total free-energy change in a thermodynamic cycle is zero, the QM/MM free energy between the two states can therefore be obtained as:

$$\Delta\Delta G_{0\to 1}^{\mathrm{QM/MM}} = \Delta G_{0}^{\mathrm{MM}\to\mathrm{QM}} + \Delta G_{0\to 1}^{\mathrm{MM}} - \Delta G_{1}^{\mathrm{MM}\to\mathrm{QM}} \qquad (5.1)$$

5.2 Reference-potential methods and statistical estimators

The term $\Delta\Delta G_{0\to1}^{\mathrm{MM}\to\mathrm{QM}}$ can be estimated in many different ways, e.g. using any of the FEP methods described above (EA, TI or BAR). However, most of these require sampling at the QM/MM level, which is very expensive. This may be avoided by using EA, based only on MM sampling at the endpoint, i.e. by:

$$\Delta G_{s,L_i}^{\mathrm{MM}\to\mathrm{QM}} = -RT\ln\left\langle e^{-\left[E_{L_i}^{\mathrm{QM}}-E_{L_i}^{\mathrm{MM}}\right]/RT}\right\rangle_{s,L_i} \qquad (5.2)$$

where the subscript *s* refers to the bound and free states, respectively. The challenge is that the perturbation needs to converge in a single step (i.e. it cannot be divided into many small steps with a λ coupling parameter, as is normally done for FEPs at the MM level). Therefore, we call this approach single-step EA (ssEA). It was evaluated in Paper I.

The convergence of ssEA can be improved by using the cumulant expansion to second order (described in section 4.1; ssEAc). The free energy is then obtained as:

$$\Delta G_{\mathrm{ssEAc}}^{\mathrm{QM}\to\mathrm{MM}} = \mu - \frac{\beta}{2}\sigma^2 \qquad (5.3)$$

where μ and σ^2 are the mean and variance of the $E^{\mathrm{QM}} - E^{\mathrm{MM}}$ energy distribution. This approach works best if $\Delta E = E^{\mathrm{QM}} - E^{\mathrm{MM}}$ follows a Gaussian distribution, which often is the case, as is illustrated in Figure 5.2. The ssEAc method was also employed in Paper I.

Another method to estimate the MM→QM/MM free energy without performing QM/MM sampling is the non-Boltzmann Bennett acceptance ratio approach (NBB)[51,52]. It reweights the snapshots sampled at the MM level with the QM/MM potential:

$$\Delta G_{0\to1}^{\mathrm{QM/MM}} = RT\left(\frac{\left\langle f(E_0^{\mathrm{QM}} - E_1^{\mathrm{QM}} + C)e^{E_1^{\mathrm{bias}}/RT}\right\rangle_1 \left\langle e^{E_0^{\mathrm{bias}}/RT}\right\rangle_0}{\left\langle f(E_1^{\mathrm{QM}} - E_0^{\mathrm{QM}} - C)e^{E_0^{\mathrm{bias}}/RT}\right\rangle_0 \left\langle e^{E_1^{\mathrm{bias}}/RT}\right\rangle_1}\right) + C \qquad (5.4)$$

where E_0^{bias} and E_1^{bias} and are the bias potentials between the MM and QM/MM levels of theory at the endpoints, respectively. In variance to ssEA, it requires QM/MM calculations for at least two λ-values and therefore requires twice the computational effort. On the other hand, BAR has better convergence properties than EA, especially when the overlap of the energy distributions is low. Typically, NBB is used only for the two endpoints of the perturbation (λ = 0 and 1), and therefore require QM/MM calculations at four λ-values, illustrated in Figure 5.3 and referred to as NBB4.

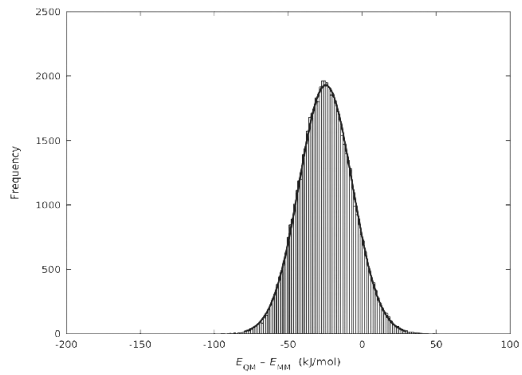


Figure 5.2. Normal distribution fit of $E^{\mathrm{QM}} - E^{\mathrm{MM}}$ from Paper I.

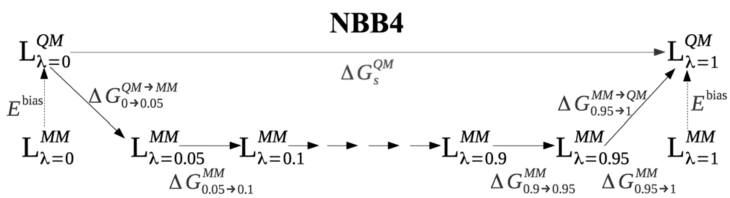


Figure 5.3. Thermodynamic cycles for NBB4 employed in Paper I.

If the QM/MM and MM distributions are too dissimilar so that ssEA or NBB converges too poorly, it may be necessary to divide also the MM→QM/MM perturbation into several steps. In analogy with the MM case, this can be done by introducing a coupling parameter, Λ , connecting the two energy functions:

$$E(\Lambda) = \Lambda\, E_{\mathrm{QM/MM}} + (1-\Lambda)\, E_{\mathrm{MM}} \qquad (5.5)$$

(this coupling parameter is distinct from λ used in the MM perturbation). In this case, sampling with the expensive QM/MM energy function is needed. We call this approach reference-potential with QM/MM sampling (RPQS), see Figure 5.4. It is in principle similar to the paradynamics approach, suggested by Warshel and coworkers[53], but paradynamics typically employs only two Λ-values and the linear-response approximation. The method was developed in Paper II and was shown to give the same results as direct QM/MM-FEP. In Paper III, we tried to speed up the approach by employing an ensemble of multiple short simulations (RPQS-MSS).

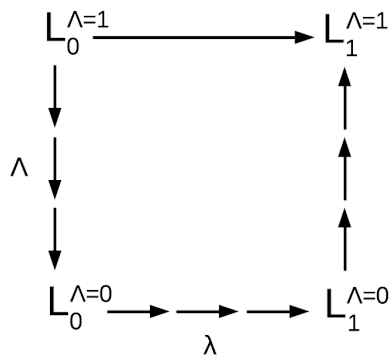


Figure 5.4. The RPQS thermodynamic cycle.

6 Binding free energies from QM/MM-minimised structures

An alternative method to obtain QM/MM free energies is to use QM-optimised structures from MM snapshots. Such an approach was originally developed by Stefan Grimme[54]. In that method, the binding free energy is determined from:

$$\Delta G_{\text{tot}} = \Delta E_{\text{QM}} + \Delta E_{\text{disp}} + \Delta G_{\text{solv}} + \Delta G_{\text{therm}} \qquad (6.1)$$

where ΔE_{QM} is the QM interaction energy, obtained with a large basis set (i.e. sufficiently large to minimise the basis-set superposition error), ΔE_{disp} is the dispersion energy, calculated with the DFT-D3 approach, third-order terms and Becke–Johnson damping, ΔG_{solv} is the COSMO-RS solvation free energy and ΔG_{therm} is a thermostatistical correction, calculated from vibrational frequencies using an ideal-gas rigid-rotor harmonic-oscillator approach. This approach has so far been employed only for host–guest systems. Then, the binding free energy is estimated from the free energies of the complex, the host and the guest molecules as:

$$\Delta G_{\text{bind}} = \Delta G_{\text{tot}}(\text{complex}) - \Delta G_{\text{tot}}(\text{host}) - \Delta G_{\text{tot}}(\text{guest}) \qquad (6.2)$$

This approach was employed in Paper IV, based on sampling with molecular dynamics simulations to obtain MM snapshots and thereafter performing QM calculations both at the semiempirical PM6-DH+ and the DFT TPSS/def2-QZVP levels of theory.

7 Main results of the thesis

The papers included in this thesis are organized as follows:

- Method development for QM/MM-FEP (Papers I, II, III, IV)
- Application of other QM/MM methods to estimate ligand-binding affinities (Papers IV).
- Method development for MM FEP (Papers V).
- Application of binding entropy calculation in the study of protein–ligand binding free energies (Papers VI).

In the following, I will discuss the main results of each paper.

Paper I. Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level.

In two earlier studies in our research group, QM/MM-FEP was attempted with the ssEA and NBB approaches, but no converged results could be obtained[55,56]. In my first paper, the convergence of QM/MM-FEP was studied for host–guest binding free energies. We showed that QM/MM-FEP can be converged to 1 kJ/mol by using 700 000 QM calculations. The precision of the method follows the expected $1/\sqrt{N}$ dependence, where N is the number of snapshots (Figure 7.1). We compared several methods to calculate the free energy, viz. NBB4, ssEA and ssEAc, and showed that ssEAc was the most cost-efficient method. However, the accuracy of the ssEAc method was not better than MM FEP (Figure 7.2).

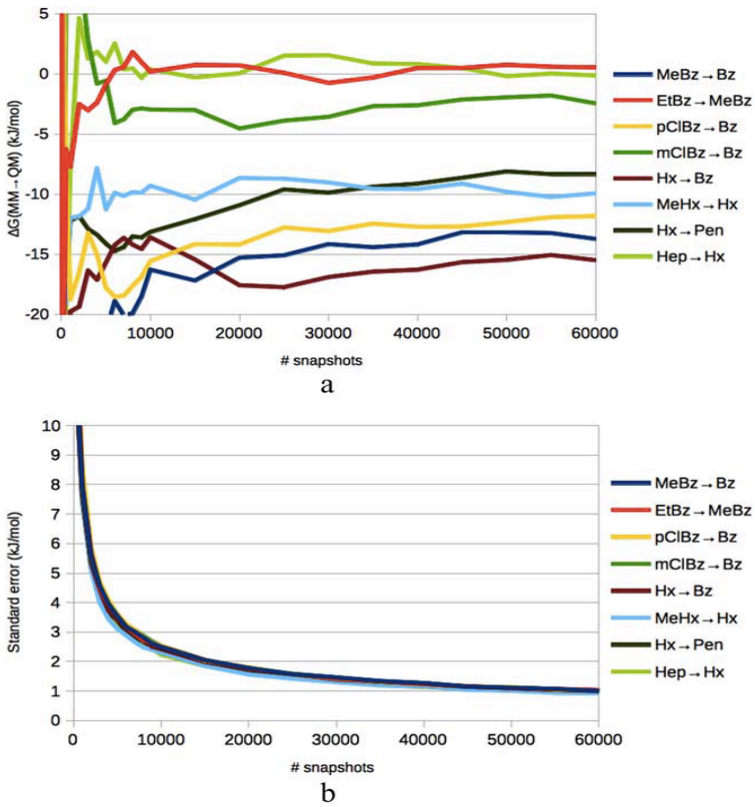


Figure 7.1. (a) Convergence of the ssEAc predictions of $\Delta\Delta G^{MM \rightarrow QM}$ with respect to the number of considered snapshots for the eight transformations (b) standard error of the calculations based on 1000 bootstraps.

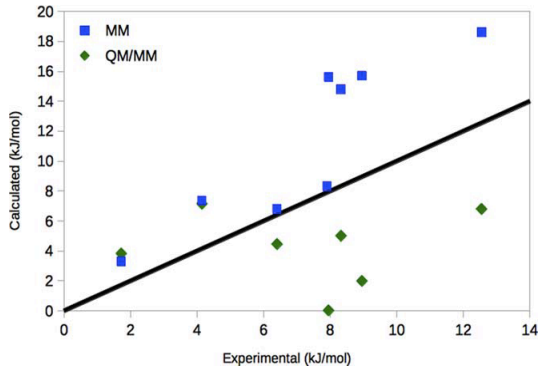


Figure 7.2. Comparison of the MM and SQM/MM energies using the ssEAc method compared to the experimental relative affinities for the eight transformations. The black line shows the perfect correlation.

Paper II. Comparison of methods to obtain ligand-binding free energies with QM/MM methods.

In the second paper, the convergence of the QM/MM energies are studied when explicit QM/MM MD simulations are performed and the MM→QM/MM perturbation is subdivided into several steps with another coupling parameter Λ in QM/MM method space, i.e. the RPQS method. The paper is the first to explicitly verify that the reference-potential method is in agreement with direct alchemical QM/MM free-energy perturbation and that the two methods give identical results. Furthermore, this paper demonstrated that the reference-potential method has a four times lower computational cost than the direct QM/MM-FEP (Table 7.1, Figure 7.3).

Table 7.1. Comparison of the reference-potential method and the direct QM/MM-FEP sampling. Unit: kJ/mol.

	RPQS 4 + 4 Λ	**Direct QM/MM-FEP 17-18 λ**	**Difference**
pClBz → Bz	23.3 ± 0.5	23.2 ± 0.9	0.1
mClBz → Bz	9.4 ± 0.5	11.3 ± 0.8	1.9

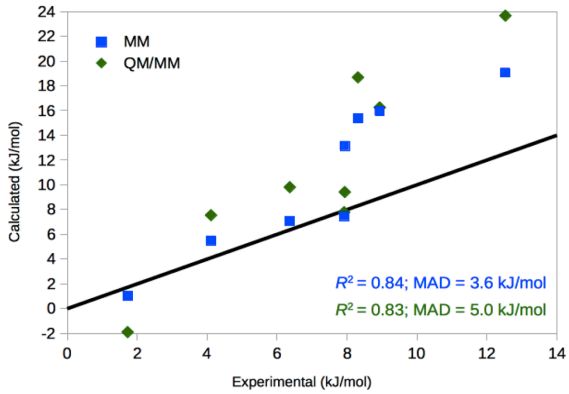


Figure 7.3. Comparison of the experimental and calculated affinities obtained with either MM or RPQS with nine Λ-values. The line shows the perfect correlation.

Paper III. Relative ligand-binding free energies calculated from multiple short QM/MM MD Simulations.

In Paper III, the computational efficiency of the RPQS approach developed in Paper II was investigated and it was examined whether it could be sped up by using multiple short FEP calculations, employing the fact that the MM simulations already thoroughly sample the phase space (the RPQS-MSS method). The paper shows that eight free-energy perturbations for the octa-acid guest molecules converges to within 1 kJ/mol in less than 50 ps of sampling using an ensemble of 100 independent simulations per perturbation (Figure 7.4). For the ninth ligand (Figure 7.5), longer simulations were needed (~70 ps), owing to a mismatch between the preferred structures of the MM and QM energy functions.

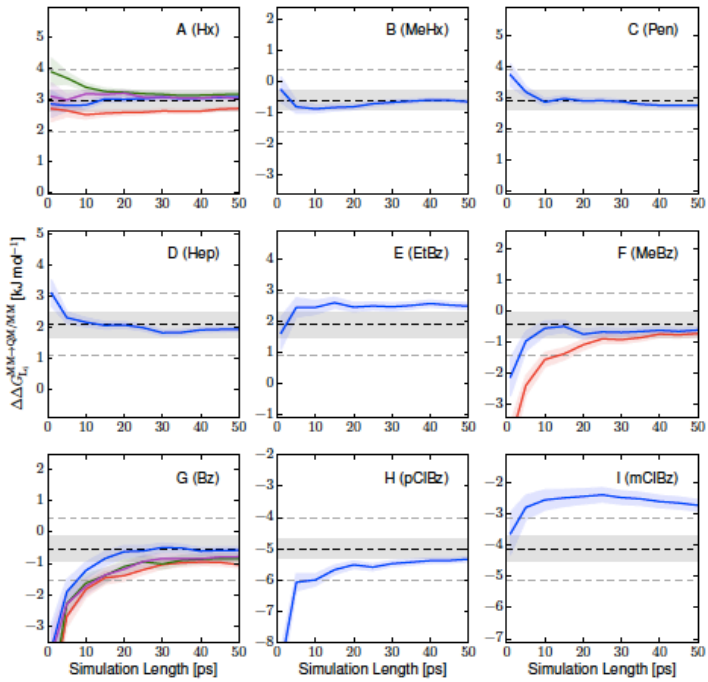


Figure 7.4. Convergence profiles for the nine ligands in this study as a function of simulation time per window.

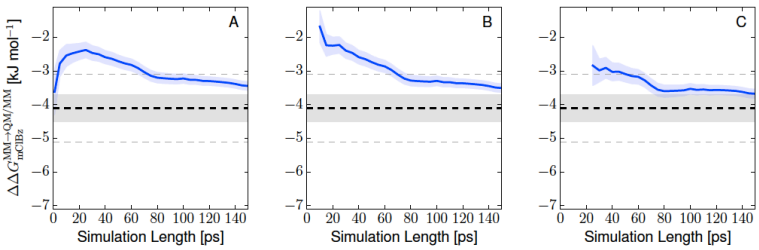
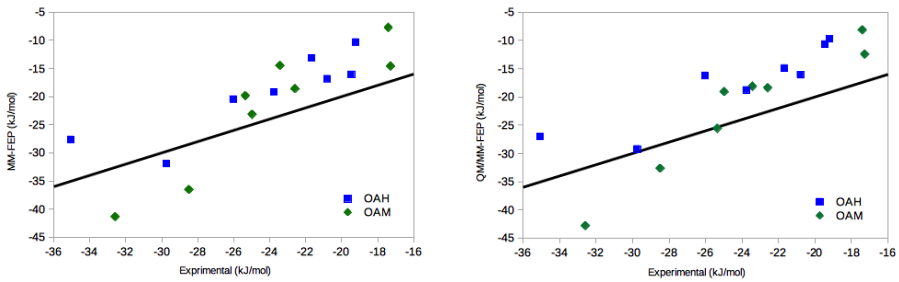


Figure 7.5. Convergence profiles for the ninth ligand in this study as a function of simulation time per window for a 150 ps trajectory.

Paper IV. Binding free energies in the SAMPL6 octa-acid host–guest challenge calculated with MM and QM methods.

In Paper IV, we used a standard MM-FEP protocol[34], RPQS and binding free energies from QM/MM-minimised structures[54,57–59] at the PM6-DH+ and TPSS-D3 levels of theory for the blind-prediction challenge SAMPL6. The best method was found to be the RPQS method (Paper II), which gave a MAD of 2.4–5.0 kJ/mol, r^2 = 0.81–0.93 and $\tau_{r,90}$ = 0.84–1.00. It was (together with standard MM FEP) one of the best five methods in the competition (Figure 7.6).



a **b**

Figure 7.6. Comparison of the experimental and calculated absolute affinities obtained with the (a) MM-FEP and (b) QM/MM-FEP methods. The black line shows the perfect correlation. OAH and OAM are two variants of the octa-acid deep-cavity host.

Paper V. Binding affinities of the farnesoid X receptor in the D3R Grand Challenge 2 estimated by free-energy perturbation and docking.

In this paper, the charge correction scheme by Rocklin et al.[41] 2013 was implemented together with the recent free-energy perturbation protocol in the AMBER software for blind-prediction in the D3R Grand Challenge 2. The results gave a MAD of 7.5 kJ/mol compared to the experimental estimates and a squared correlation coefficient r^2 = 0.1. The results suggested that including a charge correction for free-energy perturbation involving a change in the net charge improves the experimental agreement with experimental data significantly (~8 kJ/mol) on average and in the correct direction. These results were among the four best in the competition out of 22 submissions (Figure 7.7).

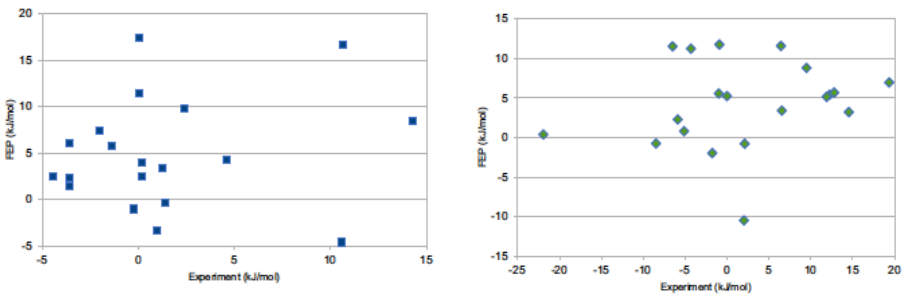


Figure 7.7. Comparison between the experimental and calculated binding free energies for the two FEP sets in the D3R Grand Challenge 3, (a) FE set 1 (b) FE set 2.

Paper VI. Detailed characterization of the binding of diastereomeric ligands to galectin-3.

Previously, binding entropy calculation has been applied to protein–ligand binding by several groups[19,60,61]. Here, we study two diastereomeric ligands binding to the protein galectin-3. We compare small structural differences in the ligands, which affect the conformational entropy of protein–ligand complexes. One of the two ligands, show two conformations in the crystal structure. We calculate entropies with the method developed by Genheden et al.[19] 2009, using a windowing scheme to calculate average conformational entropies by dihedral histogramming. The results indicate that $-T\Delta\Delta S_{conf}$ is 9 ± 5 kJ/mol between the two complexes. The calculated relative conformational entropy agrees with the experimental conformational entropy (from backbone and methyl groups) of 12 ± 8 kJ/mol both in sign and magnitude (Figure 7.8, Table 7.2).

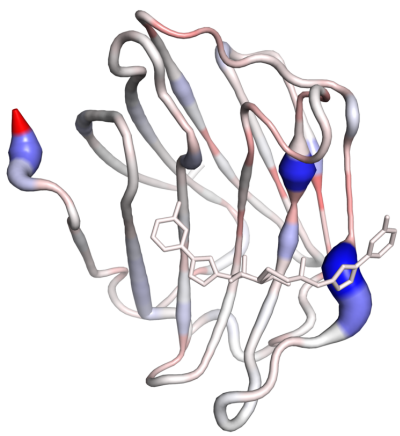


Figure 7.8. Conformational entropy contributions to ligand binding reported per residue. $T\Delta\Delta S_{conf}$ is color coded onto the galectin-3 structure with blue hues indicating positive values and red hues indicating negative values, with the color intensity ranging from weak (white) for $T\Delta\Delta S_{conf} = 0$ to intense (maximally blue or red) for $|T\Delta\Delta S_{conf}| = 3$ kJ/mol.

Table 7.2. Conformational entropy differences between the various R– and S–galectin-3 complexes and the apo protein, obtained from the MD simulations[a].

	R–apo	S–apo	S2–apo
$-T\Delta\Delta S_{conf}$ [b]	43 ± 5	33 ± 5	32 ±5
$-T\Delta\Delta S_{conf}$ [c]	67 ± 5	57 ± 4	58 ± 5

[a] S2 is the second conformation of the S ligand in complex with galectin-3.

[b] Includes all protein dihedrals.

[c] Includes all protein and ligand dihedrals.

8 Conclusions

In this thesis, the main aim of converging QM/MM free-energy perturbations for binding free energies to a precision of 1 kJ/mol has been obtained. The converged QM/MM binding free energies have a correlation with experiment of r^2 = 0.81–0.96 and MAD = 2.4–6.7 kJ/mol, whereas MM FEP gives MADs of 2.6–7.5 kJ/mol.

Below conclusions from the various papers are summarised:

I. It is possible to converge QM/MM free-energy perturbation to a precision of 1 kJ/mol, which is necessary to differentiate between ligands differing by a factor of ten in the binding constant K_{bind}.

II. It is four times more computationally efficient to use QM/MM sampling with the RPQS method than the direct alchemical approach at the QM/MM level of theory.

III. The use of multiple short simulations with RPQS speeds up the method by a factor of four.

IV. RPQS gives slightly better estimates of binding free energies for host–guest systems than FEP at the MM level and appreciably better than QM/MM minimised structures.

V. Charge corrections for periodic free-energy simulations gives significant improvements in the free energy (~8 kJ/mol) in protein–ligand binding.

VI. Windowing of trajectories is necessary to converge estimates of conformational entropy using dihedral histogramming. Conformational entropy calculations can probe small stereochemical differences of drug molecules in protein–ligand binding.

9 Outlook

Based on the results in this thesis, converged QM/MM-FEP calculations at the semiempirical level of theory can sometimes give better results than FEP at the MM level of theory. A natural extension of this method would be to use a more accurate QM method, e.g. density functional theory. The method should also be applied to enzyme reaction mechanisms. For this purpose, this thesis will provide support to carry out such evaluations. RPQS may be applied in protein–ligand binding and the relationship between the accuracy, the convergence and size of the QM system should be investigated. Furthermore, RPQS-MSS with QM drug molecules and QM binding site water structure[62] may be evaluated. An alternative to the MSS approach is to instead use non-equilibrium simulations and the Jarzynski equation[63,64].

However, the present thesis indicates that a more accurate force field does not fully improve the accuracy of FEP. RPQS and enhanced sampling methods could be combined into new methods (and correct for introduced sampling bias e.g. with FEP between biased and unbiased MM levels of theory[65] or from implicit to explicit water[66]) with the opportunity to reduce the computational cost for the conformational sampling in FEP[16]. Furthermore, evaluations of enhanced sampling in FEP should address the cycle closure hysteresis. In addition, employing soft-core potentials[67] for the full ligands in protein–ligand binding could also be assessed in further studies of the accuracy of FEP at the MM level of theory.

References

(1) Gohlke, H.; Klebe, G. *Angew. Chemie - Int. Ed.* **2002**, *41* (15), 2644–2676.

(2) Kontoyianni, M.; Madhav, P.; Seibel, E. S. *Curr. Med. Chem.* **2008**, *15* (2), 107–116.

(3) Kollman, P. A.; Massova, I.; Reyes, C. M.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M. C.; Lee T; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(4) Genheden, S.; Ryde, U. *Expert Opin. Drug Discov.* **2015**, *441* (October), 1–13.

(5) Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002**, *35*, 358–365.

(6) Wereszczynski, J.; McCammon, J. A. *Q. Rev. Biophys.* **2012**, *45* (1), 1–25.

(7) Hansen, N.; Van Gunsteren, W. F. *J. Chem. Theory Comput.* **2014**, *10* (7), 2632–2647.

(8) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136* (3B), B864–B871.

(9) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103* (2), 227–249.

(10) Karplus, M.; Levitt, M.; Warshel, A. **2013**.

(11) Shaw, D. E.; Bowers, K. J.; Chow, E.; Eastwood, M. P.; Ierardi, D. J.; Klepeis, J. L.; Kuskin, J. S.; Larson, R. H.; Lindorff-Larsen, K.; Maragakis, P.; Moraes, M. A.; Dror, R. O.; Piana, S.; Shan, Y.; Towles, B.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis - SC '09*; ACM Press: New York, New York, USA, 2009; p 1.

(12) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. . *J. Comput. Phys.* **1977**, *23* (3), 327–341.

(13) Grime, J. M. A.; Dama, J. F.; Ganser-Pornillos, B. K.; Woodward, C. L.; Jensen, G. J.; Yeager, M.; Voth, G. A. *Nat. Commun.* **2016**, *7*, 11568.

(14) Perilla, J. R.; Schulten, K. *Nat. Commun.* **2017**, *8*, 15959.

(15) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112* (16), 6127–6129.

(16) Anandakrishnan, R.; Drozdetski, A.; Walker, R. C.; Onufriev, A. V. *Biophys. J.* **2015**, *108* (5), 1153–1164.

(17) Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115* (14), 6289–6292.

(18) Karplus, M.; Kushickt, J. N. *Macromol. J. J. Polym. Sci., Polyrn. Phys. Ed. J. ADDL Phvs. J. B. J. Polyme J. J. Phys. J. Macromol.* **1972**, *17131813* (14), 585793–1317.

(19) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. *J. Biomol. NMR* **2009**, *45* (1–2), 157–169.

(20) Genheden, S.; Ryde, U. *Phys. Chem. Chem. Phys.* **2012**, *14* (24), 8662.

(21) Edholm, O.; Berendsen, H. J. C.; van der Ploeg, P. *Mol. Phys.* **1983**, *48* (2), 379–388.

(22) Prompers, J. J.; Brüschweiler, R. *J. Am. Chem. Soc.* **2002**, *124* (16), 4522–4534.

(23) Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput. Aided. Mol. Des.* **1998**, *12* (1), 27–35.

(24) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3* (5), 300–313.

(25) Bennett, C. H. *J. Comput. Phys.* **1976**, *22* (2), 245–268.

(26) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91* (14), 140601.

(27) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129* (12), 124105.

(28) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22* (8), 1420–1426.

(29) Pohorille, A. (Andrew); Chipot, C. (Christophe). *Free energy calculations : theory and applications in chemistry and biology*; Springer, 2007.

(30) Archontis, G.; Karplus, M. *J. Chem. Phys.* **1998**, *105* (24), 11246.

(31) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. .

(32) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2009**, *131* (42), 15403–15411.

(33) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Gilson, M. K. *J. Comput. Aided. Mol. Des.* **2017**, *31* (1), 133–145.

(34) Mikulskis, P.; Genheden, S.; Ryde, U. *J. Chem. Inf. Model.* **2014**, *54* (10), 2794–2806.

(35) Bhattacharyya A. *Bull. Calcutta Math. Soc* **1943**.

(36) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123* (5), 54103.

(37) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114* (32), 10235–10253.

(38) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94* (13), 138302.

(39) Li, H.; Yang, W. *Chem. Phys. Lett.* **2007**, *440* (1–3), 155–159.

(40) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. *J. Chem. Phys.* **2013**, *139* (18), 184103.

(41) Rocklin, G. J.; Boyce, S. E.; Fischer, M.; Fish, I.; Mobley, D. L.; Shoichet, B. K.; Dill, K. A. *J Mol Biol. J Mol Biol. Novemb.* **2013**, *15* (42522), 4569–4583.

(42) Ryde, U.; Söderhjelm, P. *Chem. Rev.* **2016**, *116* (9), 5520–5566.

(43) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz, K. M. *Drug Discov. Today* **2007**, *12* (17).

(44) Korth, M. *J. Chem. Theory Comput.* **2010**, *6* (12), 3808–3816.

(45) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129* (30), 9296–9297.

(46) Rathore, R. S.; Reddy, R. N.; Kondapi, A. K.; Reddanna, P.; Reddy, M. R. *Theor. Chem. Acc.* **2012**, *131* (2), 1096.

(47) Świderek, K.; Martí, S.; Moliner, V. *Phys. Chem. Chem. Phys.* **2012**, *14* (36), 12614.

(48) Muller, R. P.; Warshel, A. *J. Phys. Chem* **1995**, *99*, 17516–17524.

(49) Iftimie, R.; Salahub, D.; Wei, D.; Schofield, J. *J. Chem. Phys.* **2000**, *113* (12), 4852.

(50) Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J. *J. Chem. Phys.* **1999**, *110* (3), 1329.

(51) König, G.; Boresch, S. *J. Comput. Chem.* **2011**, *32* (6), 1082–1090.

(52) König, G.; Hudson, P. S.; Boresch, S.; Woodcock, H. L. *J. Chem. Theory Comput.* **2014**, *10* (4), 1406–1419.

(53) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B* **2011**, *115* (24), 7950–7962.

(54) Grimme, S. *Chem. - A Eur. J.* **2012**, *18* (32), 9955–9964.

(55) Hu, L.; Eliasson, J.; Heimdal, J.; Ryde, U. *J. Phys. Chem. A* **2009**, *113* (43), 11793–11800.

(56) Heimdal, J.; Ryde, U. *Phys. Chem. Chem. Phys.* **2012**, *14* (36), 12592.

(57) Antony, J.; Sure, R.; Grimme, S. *Chem. Commun.* **2015**, *51* (10), 1764–1774.

(58) Mikulskis, P.; Cioloboc, D.; Andrejić, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Söderhjelm, P.; Ryde, U. *J. Comput. Aided. Mol. Des.* **2014**, *28* (4), 375–400.

(59) Caldararu, O.; Olsson, M. A.; Riplinger, C.; Neese, F.; Ryde, U. *J. Comput. Aided. Mol. Des.* **2017**, *31* (1), 87–106.

(60) Carlsson, J.; Åqvist, J. *J. Phys. Chem. B* **2005**, *109* (13), 6448–6456.

(61) Singh, N.; Warshel, A. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (7), 1724–1735.

(62) Fox, S. J.; Pittock, C.; Tautermann, C. S.; Fox, T.; Christ, C.; Malcolm, N. O. J.; Essex, J. W.; Skylaris, C. K. *J. Phys. Chem. B* **2013**, *117* (32), 9478–9485.

(63) Jarzynski, C. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **1997**, *56* (5), 5018–5035.

(64) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78* (14), 2690–2693.

(65) König, G.; Miller, B. T.; Boresch, S.; Wu, X.; Brooks, B. R. *J. Chem. Theory Comput.* **2012**, *8* (10), 3650–3662.

(66) König, G.; Bruckner, S.; Boresch, S. *J. Comput. Chem.* **2009**, *30* (11), 1712–1718.

(67) Pitera, J. W.; van Gunsteren, W. F. *Mol. Simul.* **2002**, *28* (1–2), 45–65.

Acknowledgments

In this thesis work, I would like to thank my supervisor Ulf Ryde for teaching me scientific judgement in my work and helping out troubleshooting. I would also like to thank some people in collaborations, including Alfonso Garcia-Sosa for his patience that finally was fruitful; Casper Steinmann, Erik D. Hedegård, and Octav Caldararu for their humour; Majda M. Ignjatović and Meiting Wang for using the RPQS method. I would especially like to thank Svante Hedström and Paulius Mikulskis for their friendship when I was beginning my PhD position. I would like to thank Geng Dong for many discussions. Thanks to Per-Åke Malmqvist and Alexei Abrikossov for sharing their vast knowledge during lunches together and thanks Valera Veryazov for many funny stories. I would like to thank Lili Cao for introducing one of my favourite dishes, hot pot, to the research group. To my parents and my brothers Albin and Kristian, I would like to thank you for your patience with my higher studies. Yanzi, I met you during my PhD studies, I'm especially grateful for your support (and amazing cooking). I love you; "you are my pillow and I am your stove."

In addition, I would like to mention some people who have participated in the multidisciplinary KAW project DECREC with me; in particular Mikael Akke for his senior leadership, Maria Luisa Vertaramo for the synthesis effort of the diastereomeric ligands that Paper VI is based on, Olof Stenström for experimental entropies, and Johan Wallerstein as well as Kristoffer Peterson for conducting many breakfast meetings. I also would like to thank Hakon Leffler, Ulf J. Nilsson, Esko Oksanen, and Derek Logan for sharing funny stories during many DECREC meetings.

As some ending words, I would like to tell a brief anecdote. When I was just beginning my PhD studies, the Nobel laureate Michael Levitt (who pioneered QM/MM with Warshel and Karplus) gave his Nobel lecture at Lund University. Then, I had the opportunity to meet him in a seminar session. A dozen of privileged PhD students had a round of short presentations about their research topic. I presented that I was going to converge free-energy perturbations MM→QM to 1 kJ/mol precision. Excited about what Levitt would say as his brilliant comments, standing a meter away from him, Levitt gave me this somewhat common-sense advice: “you need to start looking at the data”. I later found use for his advice in obtaining the ssEAc method in Paper I.

Paper I

Martin A. Olsson, Pär Söderhjelm, Ulf Ryde, *Journal of Computational Chemistry,* **2016**, 37, 1589–1600

Journal of COMPUTATIONAL CHEMISTRY WWW.C-CHEM.ORG FULL PAPER

Converging Ligand-Binding Free Energies Obtained with Free-Energy Perturbations at the Quantum Mechanical Level

Martin A. Olsson,[a] Pär Söderhjelm,[b] and Ulf Ryde*[a]

In this article, the convergence of quantum mechanical (QM) free-energy simulations based on molecular dynamics simulations at the molecular mechanics (MM) level has been investigated. We have estimated relative free energies for the binding of nine cyclic carboxylate ligands to the octa-acid deep-cavity host, including the host, the ligand, and all water molecules within 4.5 Å of the ligand in the QM calculations (158–224 atoms). We use single-step exponential averaging (ssEA) and the non-Boltzmann Bennett acceptance ratio (NBB) methods to estimate QM/MM free energy with the semi-empirical PM6-DH2X method, both based on interaction energies. We show that ssEA with cumulant expansion gives a better convergence and uses half as many QM calculations as NBB, although the two methods give consistent results. With 720,000 QM calculations per transformation, QM/MM free-energy estimates with a precision of 1 kJ/mol can be obtained for all eight relative energies with ssEA, showing that this approach can be used to calculate converged QM/MM binding free energies for realistic systems and large QM partitions. © 2016 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24375

Introduction

One of the largest challenges for computational chemistry is to develop methods to estimate binding energies of small molecules to biomacromolecules. If such energies could be accurately estimated, important parts of drug development could be performed computationally. Consequently, many methods have been developed with this aim, ranging from fast scoring methods, over end-point methods, to strict free-energy simulation (FES) methods.[1–3] Owing to the size of the macromolecule, such calculations have typically been performed at the molecular-mechanics (MM) level of theory. However, it is well-known that the MM force fields used for biochemical molecules involve severe approximations, for example, omitting polarisation, higher-order multipoles, charge transfer, and charge penetration. All these effects are automatically included in quantum-mechanical (QM) calculations. Therefore, there have lately been quite some interest to improve binding-affinity calculations by QM methods,[4–6] for example, as a postprocessing of scoring calculations, improvement of docking calculations, or as a component of end-point calculations.[7–15] Many different QM methods have been employed, ranging from semiempirical QM (SQM) methods,[7,10,12] via dispersion-corrected density-functional theory (DFT) methods,[13,14] and many-body perturbation theory,[11] to coupled-cluster methods.[13,15] Some calculations involved only the ligand in the QM calculations,[8,9] whereas other included also the near-by groups,[11,13–15] or even the whole system.[7,10,12]

It would be even better if QM calculations could be combined with the FES methods, which in principle should give correct results, if used with a perfect energy function and complete sampling of all relevant parts of the phase space. Unfortunately, QM methods are extremely demanding in terms of computational time and memory requirements. Currently, QM energy calculations can be performed for a full protein at the SQM level, whereas more accurate DFT calculations can be performed on one or a few thousands of atoms, and very accurate high-level QM calculations, such as the gold-standard CCSD(T) method can only be applied to a few tens of atoms. Moreover, FES methods are based on extensive sampling of the phase space, typically involving 10^7-10^8 energy calculations in a molecular dynamics or Monte Carlo simulation. Therefore, some sort of approximation is needed to perform FES calculations at the QM level. One approach is to use QM for only a small, but interesting, part of the system (e.g., the ligand) and MM for the remainder, the QM/MM approach. A few full FES ligand-binding studies have been published with such a partitioning, treating only the ligand by QM and using SQM calculations.[16–18]

Another approach is to perform the sampling at the MM level and then evaluate QM/MM energies only for a restricted number

[a] *M. A. Olsson, U. Ryde*
Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, Lund, SE-221 00, Sweden
E-mail: Ulf.Ryde@teokem.lu.se

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

Contract grant sponsor: Swedish research council (project 2014-5540); Contract grant sponsor: Knut and Alice Wallenberg Foundation; Contract grant number: KAW 2013.0022

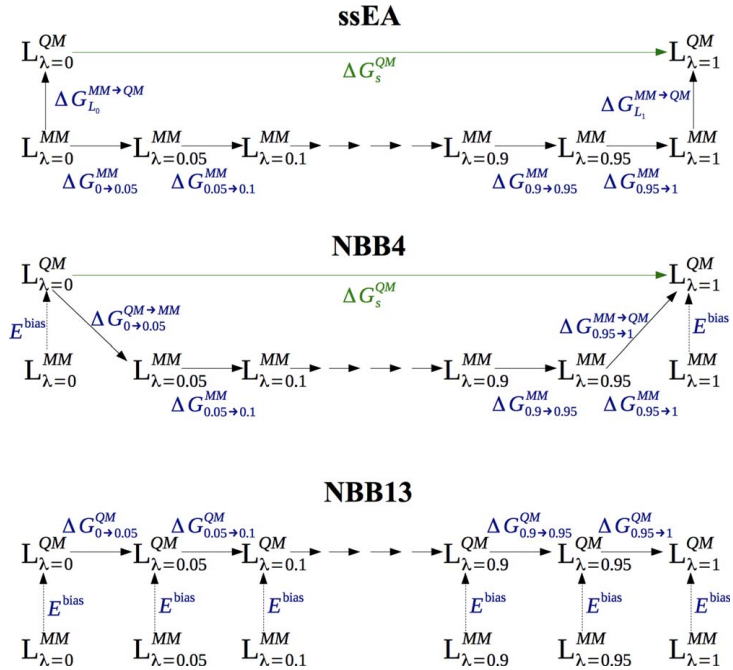


Figure 1. The various thermodynamic cycles employed in the ssEA, NBB4, and NBB13 methods to calculate binding free energies at the QM level. The cycles apply for the ligand simulated both with and without the host, giving either ΔG^{QM}_{bound} or ΔG^{QM}_{free} in eq. (2) (indicated by ΔG^{QM}_{s} in the figures). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of snapshots. Valid QM/MM free energies can be obtained either by a MM→QM/MM FES calculation, employing the thermodynamic cycle in Figure 1a,[19–21] or by reweighting of the MM snapshots toward the QM/MM energy function (Figs. 1b and 1c).[22] Such approaches have been used for ligand binding,[13,23–25] as well as for solvation free energies[26–31] and quite extensively for enzyme reactions.[19–21,32–34] The challenge with this approach is to obtain converged results for the MM→QM/MM perturbation, which must be performed in a single step to avoid the need of QM/MM sampling, that is, to ensure that the overlap of the MM and QM/MM potentials is large enough (a few approaches involving QM/MM sampling have been suggested[24,26,34–36]). For enzyme reactions, proper convergence has been obtained by keeping the QM system fixed;[19–21] without this approximation, very poor convergence has been observed, which could only partly be decreased by employing SQM/MM sampling.[34] For binding affinities, such an approximation seems inappropriate, as the entropy and reorganisation of the ligand is expected to be important for the binding.

Essex and coworkers have addressed this problem by considering only the electronic polarisation energy, which seems to give converged single-step MM→QM/MM energies calculated by exponential averaging (ssEA; i.e., using the Zwanzig free-energy perturbation approach;[37] Fig. 1a) with ~24,000 QM calculations for flexible ligands bound to cyclooxygenase-2, as well as for small molecules in water solution, in both cases with only the ligand treated by QM.[23,28] However, they have also obtained converged QM/MM solvation free energies for small phenol analogues, including 200 water molecules in the QM calculations, considering interaction energies with only 1080 QM calculations.[27] By performing full QM simulations, they have also shown that interaction energies (in contrast to total QM energies) give converged and consistent free energies for the MM→QM perturbation.[38]

König et al. instead reweighted the MM snapshots with QM energies, using the non-Boltzmann Bennett acceptance ratio method (NBB; Fig. 1b).[22] With this approach, they have obtained converged QM/MM hydration free energies using 4000–60,000 QM calculations, treating only the ligand by QM.[29,30]

On the other hand, Mulholland and coworkers used full QM/MM Monte Carlo simulations, but employed the Metropolis–Hastings approach to reduce the number of QM calculations required.[26] They have studied the relative hydration energy of water and methanol, as well as the binding of water molecules to neuraminidase, treating only the ligand by QM.[24,39] Still, the approach is very demanding, requiring $1.2–1.6\cdot10^5$ QM calculations. However, recently Skylaris and coworkers have used a similar approach to calculate hydration free energies with full QM calculations, using QM/MM structures obtained by hybrid Monte Carlo simulation from MD simulations as an intermediate

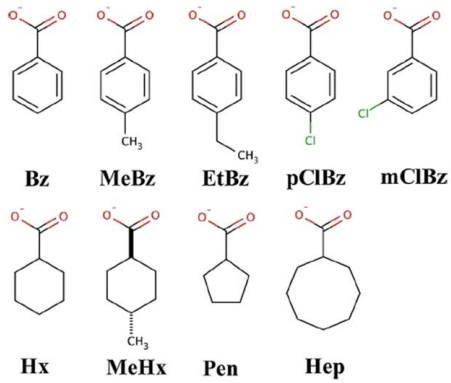


Figure 2. Guest molecules for the estimation of binding free energies to a truncated octa-acid host. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

stepping stone.[31] They obtained converged relative solvation energies by only 6000 QM calculations for each state.

We have employed both the ssEA and NBB approaches to calculate the relative binding affinities of nine cyclic carboxylic acids to the octa-acid deep-cavity host molecule (Figs. 2 and 3a) and for two synthetic disaccharides binding to galectin-3, using the full host, all protein groups, and water molecules within 6 Å of the ligand in the QM calculations (287–312 atoms for the host–guest system and 744–748 atoms for galectin-3) and dispersion-corrected density-functional theory with large basis sets (quadruple or triple zeta quality, respectively).[13,25] Unfortunately, it was not possible to obtain converged MM→QM/MM free energies for either system using 3600 QM calculations for each transformation.

The full advantage of using QM calculations is not obtained until both the ligand and at least the closest groups of the receptor (4.5–6 Å[40–42]) are included in the QM calculations. So far, no converged QM/MM binding affinities have been obtained with such an approach, owing to the use of too demanding QM methods.[13,25] Therefore, we in this article turn to the cheaper (but more approximate) SQM methods and study what is needed to obtain converged MM→QM/MM free energies for the octa-acid host–guest system. The emphasis is on convergence and what method gives the best convergence (we compare different variants of the ssEA and NBB methods), not on reproducing experimental data. We show that 720,000 QM calculations per transformation are required to converge the MM→QM free energies to within 1 kJ/mol.

Methods

Simulated system

In this article, we study the binding of nine cyclic carboxylate ligands to the octa-acid host, using experimental data from the SAMPL4 challenge.[43,44] The ligands are shown in Figure 2 and the octa-acid host in Figure 3a. Starting structures for the calculations were taken from our previous study of this system.[13] To reduce the size and the large negative charge of the host and reduce its flexibility, we deleted the four propionate groups and also the four carboxylate groups on the rim of the ring system, giving rise to a neutral cavitand (NOA) with 144 atoms, shown in Figure 3b. We will show below that this truncation has only a minor effect on relative binding affinities estimated at the MM level, but it improves the convergence of the FES calculations.

The general Amber force field[45] was used for both the NOA host and the ligands,[13] and the TIP3P force field was used for water molecules.[46] Restrained electrostatic potential (RESP) charges[47] for the ligands were taken from our previous study[13] and those of NOA were estimated in the same way: The host was optimized at the AM1 level[48] and the electrostatic potential was calculated at the Hartree–Fock/6-31G* level at points sampled around the molecule according to the Merz–Kollman scheme,[49] albeit at a higher-than-default density (10 layers with 17 points per unit area, giving ∼2000 points per atom), using the Gaussian 09 software.[50] The charges were then fitted to these potentials using the antechamber program in the Amber 14 suite.[51] It was ensured that all symmetry-equivalent atoms had the same charges (giving only 16 unique charges). The force field used for NOA is included in the Supporting Information, Table S1.

FES calculations at the MM level

All molecular dynamics (MD) simulations and FES calculations were performed with the Amber 13 (pre-release) and 14

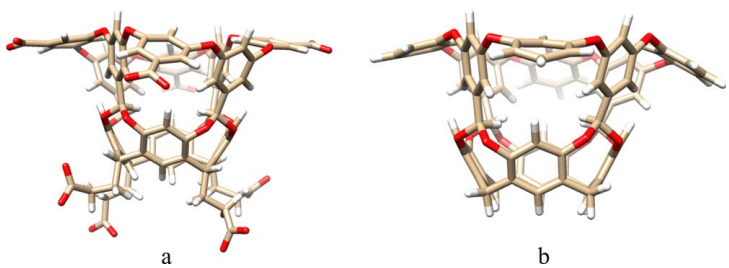


Figure 3. Structure of the full octa-acid host (a) and the neutralized host, NOA without the propionate and carboxylate groups (b).

softwares.[51] NOA and the ligands were solvated in a truncated octahedral box of water molecules, extending at least 9 Å from the solute using the leap program in the Amber suite, giving ~4100 and ~1800 atoms in total for the calculations with and without the host, respectively. Fifteen independent simulations were run for each ligand by solvating the systems in 15 different TIP3P water boxes of explicit water molecules and employing different random seeds for the starting velocities, to increase the difference between the independent simulations[52]). No counter-ions were used in the calculations (implying that a neutralising plasma were added to the systems in the simulations), because we have previously shown that they only have a minor influence on the calculated free-energy differences.[13]

The relative binding free energy between two ligands, L_0 and L_1 ($\Delta\Delta G_{bind}$), was calculated for eight transformations: MeBz→Bz, EtBz→MeBz, pClBz→Bz, mClBz→Bz, Hx→Bz, MeHx→Hx, Hx→Pen, and Hep→Hx (the names of the ligands are defined in Fig. 2). The FES calculations were run with the pmemd module of Amber,[51,53] using the dual topology scheme with both ligands in the topology files. We employed 13 states with $\lambda = 0.00, 0.05, 0.1, 0.2, \ldots, 0.8, 0.9, 0.95$, and 1.00, using a linear transformation of the potentials:

$$V_\lambda = (1-\lambda)V_0 + \lambda V_1, \quad (1)$$

where V_0 is the potential of the larger ligand and V_1 is the potential of the smaller ligand. Electrostatic and van der Waals interactions were perturbed concomitantly, using soft-core potentials for both types of interactions.[54,55] The soft-core potentials were used only for atoms differing between the two guest molecules, that is, for the transformed CH_3 →H or Cl→H groups for the MeBz→Bz, EtBz→MeBz, pClBz→Bz, mClBz→Bz, and MeHx→Hx transformations, but for all atoms in the ring system for the Hx→Bz, Hx→Pen, and Hep→Hx transformations. Test calculations have shown that using soft-core potentials for the whole guest molecule also for the smaller transformations does not change the results significantly.[13] To make the calculations comparable between the two versions of Amber, we used the keyword tishake = 1 for the Amber 14 calculations.

For each λ value, we first performed 100 steps of minimisation, with the heavy atoms of the host and guest molecules restrained toward the starting structure with a force constant of 418.4 kJ/mol/Å^2. This was followed by 20 ps constant-volume equilibration with the same restraints and 2 ns constant-pressure equilibration without any restraints. Finally, an 8 ns production simulation was run, during which structures were sampled every 2 ps. In the MD simulations, bonds involving hydrogen atoms were constrained with the SHAKE algorithm,[56] allowing for a time-step of 2 fs. In all simulations, the temperature was kept constant at 300 K using Langevin dynamics[57] with a collision frequency of 2 ps^{-1}, and the pressure was kept constant at 1 atm using a weak-coupling isotropic algorithm[58] with a relaxation time of 1 ps. Long-range electrostatics were handled by particle-mesh Ewald (PME) summation[59] with a fourth-order B spline interpolation and a tolerance of 10^{-5}. The cut-off for Lennard–Jones interactions was set to 8 Å.

The relative binding free energies were estimated using a thermodynamic cycle that relates $\Delta\Delta G_{bind}$ to the free energy of alchemically transforming L_0 into L_1 when they are either bound to the host, ΔG_{bound}, or are free in solution, ΔG_{free}[60]

$$\Delta\Delta G_{bind} = \Delta G_{bind}(L_1) - \Delta G_{bind}(L_0) = \Delta G_{bound} - \Delta G_{free} \quad (2)$$

ΔG_{bound} and ΔG_{free} can be estimated by the Bennett acceptance-ratio method[61,62] (BAR). In this approach, an MD simulation is run for each λ, with the potential in eq. (1). For each pair of neighboring λ values, A and B, the free energy difference between the two states is estimated from

$$\Delta G^{A\to B} = RT\left(\ln\frac{\langle f(V_A - V_B + C)\rangle_B}{\langle f(V_B - V_A - C)\rangle_A}\right) + C \quad (3)$$

where $f(x) = (1 + \exp(x/RT))^{-1}$ is the Fermi function, R is the gas constant, T is the temperature (which was 300 K throughout this article), and C is a constant [if the number of samples are different in the two simulations, $n_A \neq n_B$, a correction factor $\ln(n_A/n_B)$ should be added to the right-hand side of eq. (3)]. An iterative procedure is applied to find a value of C that makes the first term of the right-hand side of eq. (3) vanish. Free energies were also calculated by multi-state BAR (MBAR),[63] thermodynamic integration,[64] and exponential averaging,[37] using the pymbar software.[63] Presented results were obtained with MBAR.

SQM calculations

SQM single-point calculations were run on each of the MM snapshots, both for the simulations with and without NOA. For these calculations, water molecules were wrapped back into the original periodic box, centred on the ligand with the ptraj module. In the SQM calculations, the 48 water molecules closest to the ligand were included in the calculations without NOA, whereas the 19 water molecules closest to the C atom in the carboxylate group were included for the calculations with the ligand in NOA (in total 158–167 or 215–224 atoms, respectively; Fig. 4). This represents all water molecules within ~4.5 Å of the ligand and they were obtained in the same way as in our previous study.[13]

The PM6-DH2X method[65] was employed for the SQM calculations, that is, including dispersion, hydrogen-bond, and halogen corrections,[66–68] using the MOPAC software[69] (this was the most accurate SQM method in this software when this investigation was started). The calculations employed the keyword Precise, to enhance the energy convergence criterion to $4.2\cdot10^{-6}$ kJ/mol. For each snapshot, interaction energies were obtained by separate calculations for the complex, the guest, and the remainder (i.e., water molecules with or without NOA):[13,39]

$$\Delta E_{interact} = E_{complex} - E_{guest} - E_{remainder} \quad (4)$$

MM→QM free energies

Several different methods were tested to calculate the MM→QM free energies. First, the QM interaction energies

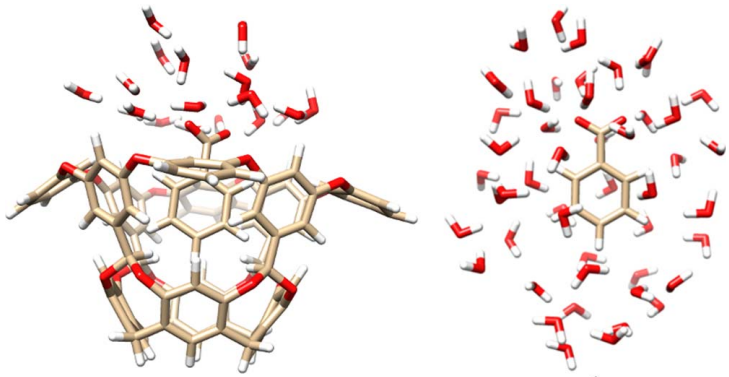


Figure 4. Example of structures used for the PM6-DH2X calculations, including 19 or 48 water molecules for the calculations with (a) and without (b) NOA, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

were used directly to calculate binding free energies for all λ values with the MBAR approach, that is, ignoring the fact that the MD simulations were not performed at the QM level. This will be called the QM-MBAR approach.

Second, we employed ssEA calculations.[13,19,20,23,27,28] In these, we employ the thermodynamic cycle in Figure 1a, showing that $\Delta\Delta G_{\text{bind}}$ is first estimated at the MM level and then two single-step FEP calculations are used to calculate the effect of changing the energy function from MM to QM, one for each of the two ligands:

$$\Delta G_s^{\text{QM}} = \Delta G_s^{\text{MM}} + \Delta G_{s,L_1}^{\text{MM}\to\text{QM}} - \Delta G_{s,L_0}^{\text{MM}\to\text{QM}} \tag{5}$$

where ΔG_s^{MM} is the free energy of the transformation at the MM level for either the bound or free states [subscript s; i.e., ΔG_{bound} or ΔG_{free} in eq. (2)], obtained by the standard MBAR approach, and the other two terms are correction terms for going from the MM potential to the QM potential. The latter corrections need to be evaluated only at the endpoints of the transformation, that is, for L_0 in the $\lambda = 0.00$ snapshots and for L_1 in the $\lambda = 1.00$ snapshots (for both the bound and free simulations). Each correction was evaluated either using exponential averaging (ssEA)[37]:

$$\Delta G_{s,L_i}^{\text{MM}\to\text{QM}} = -RT\ln\left\langle \exp\left(-\left[E_{L_i}^{\text{QM}} - E_{L_i}^{\text{MM}}\right]/RT\right)\right\rangle_{s,L_i} \tag{6}$$

or by using the cumulant expansion to the second order ($\Delta G = \mu - \frac{\sigma^2}{2RT}$, where μ is the average and σ the standard deviation of the $E_{L_i}^{\text{QM}} - E_{L_i}^{\text{MM}}$ distribution; ssEAc),[70,71] which is exact if the energy differences follow a Gaussian distribution.

Third, we employed the NBB approach to reweight the snapshots.[22] This method evaluates the free energy according to:

$$\Delta G^{\text{A}\to\text{B}} = RT\left(\frac{\left\langle f\left(E_{\text{A}}^{\text{QM}} - E_{\text{B}}^{\text{QM}} + C\right)\exp\left(E_{\text{B}}^{\text{bias}}/RT\right)\right\rangle_{\text{B}}\left\langle \exp\left(E_{\text{A}}^{\text{bias}}/RT\right)\right\rangle_{\text{A}}}{\left\langle f\left(E_{\text{B}}^{\text{QM}} - E_{\text{A}}^{\text{QM}} - C\right)\exp\left(E_{\text{A}}^{\text{bias}}/RT\right)\right\rangle_{\text{A}}\left\langle \exp\left(E_{\text{B}}^{\text{bias}}/RT\right)\right\rangle_{\text{B}}}\right) + C \tag{7}$$

where $E^{\text{bias}} = E^{\text{MM}} - E^{\text{QM}}$. This bias is a correction for the fact that the simulations are performed at the MM level, but the energies are calculated at the QM level. The advantage with NBB is that the free energies are calculated with BAR, which has better convergence properties than EA, especially when the overlap is poor.[62] The disadvantage is that at least twice as many QM calculations are needed, because BAR improves the convergence by employing the information from both a forward and backward calculation. Two different approaches to obtain the net binding free energies were used, as is illustrated in Figure 1. In the first, QM energies were calculated for all 13 λ values in the perturbation (Fig. 1c). This approach will be called NBB13 in the following. In the second approach, NBB was used only for the first two and last two λ values in the perturbation (Fig. 1b), as has been suggested by König and coworkers.[29,30] Thus, the net binding energy was obtained from

$$\begin{aligned}\Delta G_s^{\text{QM}} = {} & \Delta G_s^{\text{QM}(\lambda=0)\to\text{MM}(\lambda=0.05)} + \Delta G_s^{\text{MM}(\lambda=0.05\to\lambda=0.95)} \\ & + \Delta G_s^{\text{MM}(\lambda=0.95)\to\text{QM}(\lambda=1)}\end{aligned} \tag{8}$$

and the MM→QM energies were obtained from

$$\begin{aligned}&\Delta G_s^{\text{QM}(\lambda=0)\to\text{MM}(\lambda=0.05)} \\ &= RT\left(\frac{\left\langle f\left(E_{\lambda=0}^{\text{QM}} - E_{\lambda=0.05}^{\text{MM}} + C\right)\right\rangle_{s,\lambda=0.05}\left\langle \exp\left(E_{\lambda=0}^{\text{bias}}/RT\right)\right\rangle_{s,\lambda=0}}{\left\langle f\left(E_{\lambda=0.05}^{\text{MM}} - E_{\lambda=0}^{\text{QM}} - C\right)\exp\left(E_{\lambda=0}^{\text{bias}}/RT\right)\right\rangle_{s,\lambda=0}}\right) + C,\end{aligned} \tag{9}$$

because the MM($\lambda = 0.05$)→QM($\lambda = 0$) perturbation is based on the MM($\lambda = 0.05$) simulations, which are not biased, whereas the reverse transformation (QM($\lambda = 0$)→MM($\lambda = 0.05$)) is based on the MM($\lambda = 0$) simulations, rather than the correct QM($\lambda = 0$) simulations. It can be seen that QM calculations are needed only for L_0, but not for L_1. A similar equation applies for $\Delta G_s^{\text{MM}(\lambda=0.95)\to\text{QM}(\lambda=1)}$ (for which QM calculations are needed

for L_1, but not for L_0). This approach will be called NBB4.

All potential energies (E^{QM}, E^{MM}, and E^{bias}) in eqs. (6), (7), and (9) [and also eqs. (10)–(12) below] were approximated with the corresponding interaction energies, calculated by eq. (4). Moreover, the QM potential energies in the equations were calculated either for the isolated QM system (x_{QM}; i.e., the isolated guest with 49 water molecules or the host–guest complex with 19 water molecules) or for the full system with a QM/MM approach:

$$E^{QM/MM} = E^{QM}(x_{QM}) - E^{MM}(x_{QM}) + E^{MM}(x_{all}) \tag{10}$$

For the ssEA method in eqs. (5) and (6), the two approaches give the same result, because the energy difference in the exponential in eq. (6) becomes in the QM/MM case

$$\begin{aligned} E_{L_i}^{QM/MM}(x_{all}) - E_{L_i}^{MM}(x_{all}) &= E_{L_i}^{QM}(x_{QM}) - E_{L_i}^{MM}(x_{QM}) \\ + E_{L_i}^{MM}(x_{all}) - E_{L_i}^{MM}(x_{all}) &= E_{L_i}^{QM}(x_{QM}) - E_{L_i}^{MM}(x_{QM}) \end{aligned} \tag{11}$$

which is the same as in eq. (6). However, for NBB4 $E_{\lambda=0}^{bias} = E_{\lambda=0}^{MM}(x_{QM}) - E_{\lambda=0}^{QM}(x_{QM})$ remains the same according to eq. (11), but eq. (9) changes to

$$\Delta G_s^{QM/MM(\lambda=0)\to MM(\lambda=0.05)} = RT\left(\frac{\left\langle f(E_{\lambda=0}^{QM/MM} - E_{\lambda=0.05}^{MM} + C)\right\rangle_{s,\lambda=0.05} \left\langle \exp(E_{\lambda=0}^{bias}/RT)\right\rangle_{s,\lambda=0}}{\left\langle f(E_{\lambda=0.05}^{MM} - E_{\lambda=0}^{QM/MM} - C)\exp(E_{\lambda=0}^{bias}/RT)\right\rangle_{s,\lambda=0}}\right) + C \tag{12}$$

with $E^{QM/MM}$ calculated from eq. (10) and the $\Delta G_s^{MM(\lambda=0.05\to\lambda=0.95)}$ energy calculated with the full systems, periodic boundary conditions, and total energies.

Uncertainties, quality estimates, and overlap measures

Reported uncertainties are standard errors, that is, standard deviations divided by the square root of the number of samples, for example, the 15 sets of independent simulations. The uncertainties of the free-energy estimates were obtained by nonparametric bootstrap sampling (using 1000 samples) of the potential-energy differences in the BAR or NBB calculations.

The quality of the binding-affinity estimates compared to experimental data was quantified using the mean absolute deviation (MAD), the root-mean-squared deviation (RMSD), the correlation coefficient (r^2), and the slope and intercept of the best correlation line. In addition, Kendall's rank correlation coefficient was calculated for the eight transformations explicitly simulated (τ_r). The uncertainties of the quality estimates were obtained by a parametric bootstrap (using 500 samples), assuming the estimates follow a Gaussian distribution with the mean equal to the estimate and the standard deviation equal to the reported uncertainty.

To estimate the convergence of the various perturbations, six different overlap measures were employed.[72] We calculated the Bhattacharyya coefficient for the energy distribution overlap (Ω),[73] the Wu & Kofke overlap measures of the energy probability distributions (K_{AB}) and their bias metrics (Π),[74,75] the weight of the maximum term in the exponential average (w_{max}),[20] the difference of the forward and backward exponential average estimate ($\Delta\Delta G_{EA}$), and the difference between the BAR and TI estimates ($\Delta\Delta G_{TI}$, although this difference may also reflect the integration error in TI[76]).[72] We used w_{max} also to estimate the convergence of the ssEA and NBB4 calculations. In the former case, w_{max} is the weight of the maximum term in the average in eq. (6). In the latter case, w_{max} was calculated for each of the three averages in eq. (9) after convergence of C and the largest of these three values is reported. However, calculated in this way and using the same data, w_{max} for ssEA and NBB4 is identical, because the latter is always dominated by the $\left(\exp\left(E_{\lambda=0}^{bias}/RT\right)\right)$ term in eq. (9), which is the same as in eq. (6).

Result and Discussion

Binding affinities at the MM level

In this article, we study the binding of nine carboxylate ligands to the octa-acid (OA) host molecule, shown in Figures 2 and 3a. We calculate the relative binding energies of the ligands with FES methods and our goal is to obtain converged relative binding energies at the QM/MM level, without performing sampling at the QM/MM level, but including all groups within ~4.5 Å of the ligand in the QM calculations (not only the ligand as in most previous studies[23,24,28–30,39]). Our previous investigation of this system as well as the binding of two ligands to galectin-3 failed to give converged QM/MM binding energies with 3600 QM calculations at the DFT level.[13,25] Therefore, we employ here the much faster SQM PM6-DH2X method, so that we can perform enough QM calculations to ensure converged results. Moreover, we have removed the propionate and benzoate groups of the octa-acid host (yielding NOA, shown in Fig. 3b), because our previous study showed that it was hard to obtain a proper sampling of the dihedral angles of the propionate groups within a typical simulation time (4 ns).[13] Moreover, the large negative charge (−8) of the host molecule sometimes gave problems in the QM calculations.

To check that the truncation of the host does not affect the results significantly, we first calculated $\Delta\Delta G_{bind}$ for the NOA host at the MM level. From the results in Table 1, it can be seen that the calculations with NOA gave almost the same results as for the full octa-acid host[13]: For five of the transformations, the two hosts gave results that agreed within 1 kJ/mol, whereas for the remaining three transformations (EtBz→MeBz, Hx→Bz, and Hep→Hx), the results differed by 2–3 kJ/mol. However, owing to the high precision of both calculations, the difference is statistically significant for all except two of the transformations (MeBz→Bz and Hx→Pen) at the 95% level.

The results of the NOA calculations are appreciably more precise than the OA calculations (0.02–0.08, compared to 0.05–0.73 kJ/mol). This is partly owing to the longer simulations (8 ns vs. 4 ns) and the larger number of independent simulations (15 vs. 10). However, there are also clear indications that the NOA calculations are better converged than the previous calculations: The overlap measures in Table 2 show a perfect overlap for all the eight transformations with NOA with all $\Omega = 1.00$, $K_{AB} \geq 1.03$, $\Pi \geq 2.5$, $w_{max} \leq 0.03$, $\Delta\Delta G_{EA} \leq 0.08$ kJ/mol, and

Table 1. Results for the eight perturbations ($\Delta\Delta G_{bind}$ in kJ/mol) obtained at the MM level (using MBAR) for the NOA host.

Transformation	OA Exp.[44]	NOA MM calc.	OA[13] MM calc.	NOA SQM/MM
MeBz→Bz	9.0±0.5	15.71±0.02	15.94±0.05	2.0±1.3
EtBz→MeBz	1.7±0.5	3.28±0.02	1.02±0.08	3.8±1.0
pClBz→Bz	12.6±0.2	18.60±0.02	19.06±0.09	6.8±1.0
mClBz→Bz	6.4±0.3	6.81±0.03	6.11±0.12	4.5±1.0
Hx→Bz	7.9±0.4	15.60±0.05	13.14±0.32	0.0±1.1
MeHx→Hx	8.3±0.4	14.81±0.01	15.35±0.15	5.0±1.1
Hx→Pen	7.9±0.4	8.33±0.04	7.48±0.73	0.0±1.0
Hep→Hx	4.1±0.3	7.35±0.07	5.50±0.66	7.2±0.9
MAD		4.07±0.13	3.56±0.17	4.9±0.4
RMSD		4.95±0.14	4.61±0.16	5.4±0.4
r^2		0.79±0.03	0.84±0.04	0.00±0.03
slope		1.50±0.07	1.77±0.09	0.0±0.1
inter		0.41±0.60	−2.40±0.75	3.9±0.9
τ_r		1.00±0.00	1.00±0.00	1.0±0.2

For comparison, calculated (BAR)[13] and experimental[44] results obtained with the full octa-acid (OA) host are also included. For both NOA and OA, the presented calculated results are the average and standard error over the 15 or 10 independent simulations. In the last column, the SQM/MM results for the NOA host, obtained with ssEAc and 15 independent calculations are included. The six last rows give quality measures describing how well the calculations reproduce the experimental data of OA in the first column: The mean absolute deviation (MAD in kJ/mol), the root-mean-squared deviation (RMSD in kJ/mol), the correlation coefficient, the slope and intercept of the best correlation line, and Kendall's ranking correlation coefficient for only the eight considered transformations (τ_r).

$\Delta\Delta G_{TI} \leq 0.06$ kJ/mol (Ω goes from 0, no overlap to 1, perfect overlap[73]; K_{AB} goes from 0 – no overlap, via 1 – full overlap, to 2 – the first distribution is completely inside the second distribution[74,75]; a negative Π indicates poor overlap[74,75]; $1/w_{max}$ indicates how many snapshots contribute significantly to the EA estimate; $\Delta\Delta G_{EA}$ is the hysteresis in the forward and backward EA estimates; and $\Delta\Delta G_{TI}$ indicates the difference between the BAR and TI estimates). In fact, all free-energy measures estimated by PYMBAR (TI, TI_{cubic} $EA_{forward}$, $EA_{backward}$, BAR, and MBAR) agree within 0.06–0.19 kJ/mol for the eight transformations, and the most accurate BAR and MBAR results agree within 0.05 kJ/mol, indicating extremely well-converged results.

For our previous OA simulations[13] (also listed in Table 2), the convergence was appreciably worse with Ω down to 0.93, K_{AB} down to 0.79, Π down to −0.1, w_{max} up to 0.95, and $\Delta\Delta G_{EA}$ up to 16 kJ/mol, whereas $\Delta\Delta G_{TI} \leq 0.3$ kJ/mol was good. In particular, Hx→Pen and Hep→Hx transformation gave negative Π values and $w_{max} > 0.38$, which indicates that more λ values or longer simulations should have been used. This is also reflected by the larger standard error of these two estimates (0.7 kJ/mol). Moreover, the MeHx→Hx transformation gave $w_{max} = 0.95$ and $\Delta\Delta G_{EA} = 16$ kJ/mol, which indicates that the overlap was poor also for this transformation. A large part of the improvement for NOA can be attributed to the longer simulations (60,000 snapshots instead of 4000). However, if we instead consider the worst values in the 15 independent simulations of NOA, each based on 4000 snapshots, NOA still gives converged results ($\Omega = 0.99$, $K_{AB} \geq 0.93$, $\Pi \geq 1.7$, $w_{max} \leq 0.27$, $\Delta\Delta G_{EA} \leq 0.9$ kJ/mol, and $\Delta\Delta G_{TI} \leq 0.06$ kJ/mol, although both w_{max} and $\Delta\Delta G_{EA}$ have increased by a factor of 6–17). This shows that the removal of the flexible propionate groups has strongly improved the sampling for the NOA host.

Affinities at the SQM level

Next, we tried to estimate binding affinities also at the SQM/MM level using 60,000 snapshots for each λ value (in practice, we first did the calculations on 4000 snapshots and based on those results, we decided how many independent simulations were needed to converge the results to a precision of 1 kJ/mol). As detailed in the Methods section, we employed several different approaches to calculate the MM→SQM free energies. First, we tried to use the full NBB13 approach with SQM

Table 2. Overlap measures for the eight perturbations of NOA and OA, performed at the MM level, based on 60,000 (NOA) or 4000 (OA) snapshots. Each measure is the minimum (Ω, K_{AB}, and Π) or maximum (w_{max}, $\Delta\Delta G_{EA}$, and $\Delta\Delta G_{TI}$) value over the 26 λ values for the simulations with and without the host.

	Ω	K_{AB}	Π	w_{max}	$\Delta\Delta G_{EA}$	$\Delta\Delta G_{TI}$
NOA						
MeBz→Bz	1.00	1.03	2.9	0.003	0.02	0.04
EtBz→MeBz	1.00	1.04	2.7	0.004	0.03	0.03
pClBz→Bz	1.00	1.03	2.9	0.002	0.02	0.04
mClBz→Bz	1.00	1.03	2.8	0.002	0.06	0.02
Hx→Bz	1.00	1.03	2.5	0.004	0.06	0.02
MeHx→Hx	1.00	1.03	2.8	0.017	0.05	0.01
Hx→Pen	1.00	1.04	2.5	0.026	0.08	0.04
Hep→Hx	1.00	1.04	2.5	0.009	0.05	0.06
OA						
MeBz→Bz	0.99	1.01	2.3	0.010	0.20	0.13
EtBz→MeBz	1.00	1.02	2.1	0.024	0.23	0.03
pClBz→Bz	1.00	1.02	2.5	0.009	0.12	0.14
mClBz→Bz	0.99	1.02	2.2	0.019	0.24	0.02
Hx→Bz	0.99	1.00	1.2	0.106	0.94	0.21
MeHx→Hx	0.93	0.79	1.1	**0.947**	**16.19**	0.34
Hx→Pen	0.95	0.85	**0.0**	**0.716**	1.69	0.06
Hep→Hx	0.98	0.93	**−0.1**	**0.377**	**6.42**	0.03

The measures are the Bhattacharyya coefficient for the energy distribution overlap (Ω),[73] the Wu & Kofke overlap measures of the energy probability distributions (K_{AB}) and their bias metrics (Π)[74,75] the weight of the maximum term in the EA (w_{max}),[20] the difference of the forward and backward EA estimate ($\Delta\Delta G_{EA}$ in kJ/mol), and the difference between the BAR and TI estimates ($\Delta\Delta G_{TI}$ in kJ/mol). Values indicating poor overlap or bad convergence are marked in bold face ($\Omega < 0.7$, $K_{AB} < 0.7$, $\Pi < 0.5$, $w_{max} > 0.2$, $\Delta\Delta G_{EA} > 4$ kJ/mol, or $\Delta\Delta G_{TI} > 1$ kJ/mol).[72,74,75]

Table 3. NBB4, ssEA, ssEAc, and ssPA results for the eight transformations ($\Delta\Delta G_{bind}$ or $\Delta\Delta G_{MM\rightarrow QM}$ in kJ/mol).

Quantity	$\Delta\Delta G_{bind}$		$\Delta\Delta G_{MM\rightarrow QM}$				$\Delta\Delta G_{bind}$	w_{max}
Method	NBB4		ssEA	ssEAc		ssPA	ssEAc	ssEA
Averaging	all	15 indep	all	all	15 indep	all	15 indep	all
MeBz→Bz	−2.4±4.1	−1.1±2.9	−16.7±4.1	−13.8±1.0	−13.7±1.3	−11.6±0.16	3.1±1.3	0.76
EtBz→MeBz	−1.5±4.0	−4.5±3.2	−5.0±3.9	0.6±1.0	0.5±1.0	−2.8±0.17	4.4±1.0	0.87
pClBz→Bz	11.4±3.9	13.0±2.8	−10.7±4.0	−11.8±1.0	−11.8±1.0	−5.9±0.17	9.5±1.0	0.85
mClBz→Bz	31.9±7.2	9.6±3.8	11.9±4.9	−2.4±1.0	−2.4±1.0	−2.2±0.17	5.0±1.0	0.90
Hx→Bz	3.9±6.2	−3.7±2.2	−14.3±6.3	−15.5±1.0	−15.6±1.1	−29.9±0.17	3.5±1.1	0.92
MeHx→Hx	−1.4±5.1	5.6±2.9	−17.5±5.4	−10.0±0.9	−9.8±1.1	−3.8±0.16	6.8±1.1	0.96
Hx→Pen	−1.7±2.8	−2.4±3.3	−7.8±2.9	−8.4±1.0	−8.3±1.0	−4.5±0.16	0.5±1.0	0.55
Hep→Hx	17.4±2.2	11.1±2.8	9.3±2.2	−0.1±0.9	−0.2±0.9	−0.5±0.16	8.1±0.9	0.36

Results are shown for either all 60,000 snapshots in a single calculation with standard errors obtained with bootstrapping (all) or as the average over 15 independent simulations with 4000 snapshots each, obtaining standard errors from the standard deviation over these 15 results, divided by $\sqrt{15}$ (15 indep). In the last column, w_{max} is given for the ssEA–all calculation.

calculations for all λ values (Fig. 1c). However, this is very demanding, requiring $60{,}000 \cdot 13 \cdot 2 \cdot 5 = 7{,}800{,}000$ QM calculations for each transformation (60,000 snapshots, 13 λ values, two sets of simulations, that is, with or without the host, calculations with both L_0 and L_1, and three calculations for each geometry to get interaction energies from eq. (4), but $E_{remainder}$ is the same for the two ligands). Moreover, the calculations gave many numerical problems and highly uncertain results. The reason for this is partly that the MM calculations employ soft-core potentials, which increase the difference between QM and MM and therefore deteriorates the convergence of the MM→QM perturbation. Therefore, this approach was only attempted for the MeBz→Bz transformation and for 4000 snapshots, giving $\Delta\Delta G_{bind}^{NBB13} = 29 \pm 14$ kJ/mol.

We also tried to calculate the binding free energies directly with MBAR calculations based on the QM results (QM-MBAR; using the same data as NBB13), ignoring the fact that the snapshots were obtained with MD simulations at the MM, rather than the QM level.[25] However, the results based on only 4000 snapshots for the MeBz→Bz transformation was poor (23.9 ± 0.1 kJ/mol), with large differences between estimates obtained with different methods (BAR, TI, and EA) and many overlap estimates indicating poor overlap, for example, Π down to −2.0, w_{max} up to 1.0, and $\Delta\Delta G_{EA}$ up to 949 kJ/mol. Therefore, this approach was not further pursued.

Instead, we tested the NBB4 approach suggested by König and coworkers.[29,30] In this approach, NBB is used to estimate the free energy of going from $\lambda = 0.00$ with QM to $\lambda = 0.05$ with MM (and similar between $\lambda = 0.95$ and 1.00), whereas the difference between $\lambda = 0.05$ and 0.95 is estimated at the MM level, as is illustrated in Figure 1b. NBB4 requires $60{,}000 \cdot 4 \cdot 2 \cdot 3 = 1{,}440{,}000$ QM calculations (i.e., only for $\lambda = 0.00$ and 0.05 with L_0, and for $\lambda = 0.95$ and 1.00 with L_1), which is 5.4 times fewer than with NBB13.

The NBB4 results are collected in Table 3. Two sets of results are presented: The first is for an NBB4 calculation including the concatenated results of all 60,000 snapshots with the standard error estimated by bootstrapping. The second is the average over the 15 individual independent calculations with 4000 snapshots in each and the standard error calculated from the standard deviation over the 15 sets. It can be seen that the uncertainty of the former approach is somewhat larger than for the latter, 2–7 vs. 2–3 kJ/mol. The opposite is normally observed, which indicates that the results strongly depend on a few snapshots, that is, that the calculations are still poorly converged. In most cases, the results of the two sets of calculations agree within the estimated statistical uncertainty, with differences of 1–8 kJ/mol. However, for the mClBz→Bz transformation, the difference is 20 kJ/mol, showing that the NBB4 estimates do not fully show the expected statistical behavior.

Therefore, we instead tried to estimate the MM→QM free energy by the ssEA approach. A direct application of ssEA [i.e., with the full exponential averaging in eq. (6)], gave an uncertainty similar to that for NBB4, 2–6 kJ/mol (third column in Table 3). Moreover, $w_{max} = 0.36$–0.96 (last column in Table 3), showing that the exponential average is dominated by one or a few terms (snapshots).

More stable results could be obtained by using a cumulant expansion to the second order[70,71] (which is equivalent of assuming a Gaussian distribution; ssEAc). With such an approach, the uncertainty was reduced to 0.9–1.0 kJ/mol for all eight transformations (based on calculations on all 60,000 snapshots and bootstrapped uncertainties; column four in Table 3). Closely similar results were obtained by averaging the results from the 15 independent simulations (with 4000 snapshots in each; column five in Table 3): The two approaches gave results that agreed within 0.2 kJ/mol and the uncertainties (estimated from the standard deviation over the 15 simulations) also agreed within a factor of 0.9–1.3. Moreover, the average standard errors for the individual calculations based on 4000 snapshots were 3.1–4.2 times larger than the standard errors for the calculations based on 60,000 snapshots, that is, close to $\sqrt{60000/4000} = \sqrt{15} = 3.9$, following the expected $\sqrt{n}$ dependence of a normal distribution (in fact, we selected the final number of snapshots based on such an extrapolation). Figure 5 shows how the predictions converge and the precision improves with the number of snapshots. The results of the ssEA and ssEAc methods agree within 1–14 kJ/mol, which is inside the 95% confidence interval

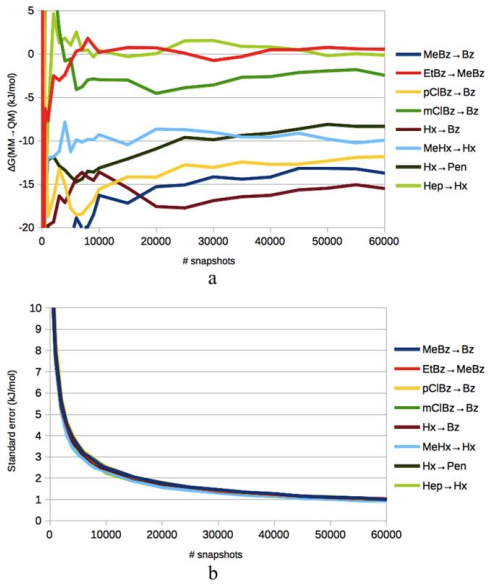


Figure 5. a) Convergence of the ssEAc predictions of $\Delta\Delta G_{MM\rightarrow QM}$ with respect to the number of considered snapshots for the eight transformations. Pane b) shows the corresponding standard error of the calculations, based on 1000 bootstraps.

(dominated by the uncertainty of ssEA) for all transformations, except mClBz→Bz and Hep→Hx, indicating that the ssEA results are not fully well-behaving. In Figure S1 in the Supporting Information, distribution and normal-probability plots are given for three of the MM→QM perturbations. It can be seen that all E^{QM}–E^{MM} distributions are very close to normal, except in the low-probability ends. The two first examples show typical results for simulations with and without the host, respectively, for which the distribution is Gaussian beyond 0.001 probability, whereas the last row shows the poorest results, for which deviations from Gaussian distribution start to emerge at 0.02 probability.

Three of the ligands are involved in more than one perturbation (Bz, MeBz, and Hx). Therefore, we have 2–4 estimates of $\Delta G_{MM\rightarrow QM}$ for these for the simulations with or without the host, and these estimates are collected in Supporting Information Table S2. In most cases, the results of these calculations agree, for example, −90.1 to −90.8 kJ/mol for Bz with the host (average -90.7 ± 0.3 kJ/mol), in agreement with the estimated uncertainty of 0.4 kJ/mol for the individual estimates. However, in three cases, one of the simulations gives deviating results by 5–9 kJ/mol. This indicates that the ssEAc results still are somewhat sensitive to rare events in the simulations and occasionally the estimated precision is too high.

Finally, we also tried to estimate the MM→QM free energy with the pure average (ssPA), instead of the exponential average in the ssEA approach. This gave well-converged results with a standard error of 0.2 kJ/mol, reflecting that a pure average has much better convergence properties than the exponential average (sixth column in Table 3). However, the pure average is an approximation to the true exponential average, an approximation that is valid only if the variation in the E^{QM}–E^{MM} energy differences is small, which is not the case. Therefore, the pure average converged to results that were different from those obtained with ssEAc. For the $\Delta\Delta G_{MM\rightarrow QM}$ corrections in Table 3, the difference between the ssPA and ssEAc results for the various transformations was up to 14 kJ/mol (5 kJ/mol on average), a difference that is statistically significant for five of the transformations. However, $\Delta\Delta G_{MM\rightarrow QM}$ is obtained as the difference of the results for the calculations with and without the NOA host, which each are the difference of the results obtained with L_0 and L_1 [eq. (2) and (5)]. For these four contributions to $\Delta\Delta G_{MM\rightarrow QM}$, the difference between ssPA and ssEAc was much larger, 59–237 kJ/mol. This clearly shows that pure averages cannot be used if you aim at an accuracy better than ∼10 kJ/mol, especially as there is no useful estimate of the true uncertainty of the approach.

The NBB and ssEA results discussed so far are not comparable, because the former are full ΔG_{bind} free energies, whereas the latter only includes the MM→QM free-energy correction ($\Delta\Delta G_{MM\rightarrow QM}$ in Table 3). Thus, to compare the results, we need to add $\Delta\Delta G^{MM}_{bind}$, obtained for the isolated QM system at the MM level. This is done in the penultimate column in Table 3. It can be seen that the NBB4 and ssEAc results agree within 1–9 kJ/mol (based on averages of the 15 independent simulations), which is reasonable, considering the quite large uncertainty of the NBB4 results.

All SQM results discussed up to this point have been based on calculations of the isolated QM system. More realistic energies can be obtained by a QM/MM approach. As discussed in the Methods section, QM and QM/MM energies give the same results for ssEA, so we easily reach a final result by adding the MM→QM ssEAc free-energy corrections in Table 3 (column five) to the $\Delta\Delta G_{bind}$ free energies, obtained at the MM level for the full periodic system from the second column in Table 1, giving the results in last column of Table 1. These results differ slightly from the results in the penultimate column in Table 3, because the latter employ MM $\Delta\Delta G^{MM}_{bind}$ binding free energies obtained for only the QM system and using interaction energies instead of total potential energies (to make the results directly comparable with the NBB4 results in Table 3 which use the same MM $\Delta\Delta G^{MM}_{bind}$ free energies). The two results differ by up to 3.5 kJ/mol (1.5 kJ/mol on average), showing that already the small QM system gives reasonable $\Delta\Delta G^{MM}_{bind}$ free-energy estimates.

From Figure 6, it can be seen that all MM→QM corrections are in the correct direction (assuming that NOA should give the same results as OA), that is, reducing the relative affinities, except in the EtBz→MeBz case, for which the MM→QM correction is close to zero. Unfortunately, the corrections are too large in six of the cases, giving too small relative binding affinities. Consequently, the SQM/MM results reproduce the experimental OA results appreciably worse than the MM results, as is shown in Table 1. For example, MAD increases to 4.9 ± 0.4 kJ/mol and r^2 vanishes. Of course, this is somewhat

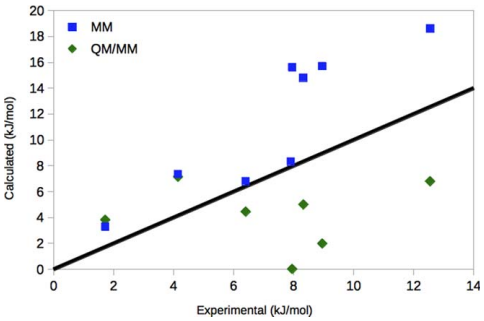


Figure 6. Comparison of the MM and SQM/MM (ssEAc) results for NOA, compared to the experimental relative affinities[44] for the eight considered transformations. The black line shows the perfect correlation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

disappointing after performing almost 6 million SQM calculations. However, this result is much better than previous attempts to obtain QM/MM FES binding free energies for the same system, which gave MADs of 17–26 kJ/mol and no convergence for the MM → QM perturbation.[13] It is also much better than approaches based on optimized structures and energies calculated by dispersion-corrected DFT methods or even CCSD(T) calculations, giving MADs of 8–37 kJ/mol.[13–15] Moreover, our aim has been to find out what is required to converge the MM→QM/MM, not to reproduce the experimental results. Therefore, we have selected a rather cheap method, PM6-D2HX, which is among the best available SQM methods, although it is appreciably worse than dispersion-corrected DFT methods.[77]

Thus, at least for this data, ssEAc with the cumulant expansion gave a lower uncertainty than NBB4. A possible reason for this is again the use of soft-core potentials in the MM calculations, which increases the difference between QM and MM: In the ssEA approach, only the $\lambda = 0.00$ and 1.00 states are considered, for which the soft-core potentials are not active. However, NBB4 considers also the $\lambda = 0.05$ and 0.95 states, for which the soft-core is active. The ssEA approach also has the advantage of using only $60{,}000 \cdot 2 \cdot 2 \cdot 3 = 720{,}000$ QM calculations (i.e., only for L_0 at $\lambda = 0.00$ and L_1 at $\lambda = 1.00$), that is, half as many as for NBB4. Moreover, the reweighting in NBB is often badly conditioned, essentially picking out a single energy (snapshot) in the second average in the nominator of eq. (9). In fact, w_{max} calculated for this method is exactly the same as for ssEA (shown in the last column in Table 3), that is, 0.36–0.96, indicating that the estimated free energies are completely dominated by one or a few QM energies (and that the use of numerous QM calculations is only a way of finding these values). This explains the rather poor convergence of both these methods. On the other hand, with the cumulant expansion, all QM values are used to estimate the average and standard deviation of the Gaussian distribution (but both values are still somewhat dominated by a few values).

All results in Table 3 were obtained using interaction energies [eq. (4)] rather than total energies. This has the advantage of making the E^{QM}–E^{MM} energy difference smaller and less varying by ignoring the difference in the two energy functions for the internal interactions within the ligand or the host. On the other hand, this is an approximation. At the MM level, it is a good approximation: For seven of the perturbations, $\Delta\Delta G_{bind}$ calculated with interaction energies (and without periodicity and Ewald summation) reproduce the results in Table 1 (based on total energies) within 0.6 kJ/mol (MAD 0.3 kJ/mol). However, for the Hx→Bz perturbation, the difference is slightly larger, 2.7 kJ/mol. On the other hand, using total energies for the MM→QM perturbation increases the variation in E^{QM}–E^{MM} by a factor of ~2, making the convergence much worse. As a consequence, $\Delta\Delta G_{MM\rightarrow QM}$ estimated by ssEAc changes by 1–12 kJ/mol (5 kJ/mol on average; pure averages change by 1–4 kJ/mol) and the bootstrapped precision estimates become very large, illustrating that these results are far from converged. This shows that interaction energies strongly improve the convergence of the MM→QM perturbations, in agreement with other studies.[38,78]

Conclusions

In this article, we have studied what is needed to obtain converged QM/MM relative binding free energies, performing sampling only at the MM level and including a significant surrounding of the ligand in the QM calculations (158–224 atoms). Previous studies with such an approach have given poorly converged results both for a host–guest system and a full protein, probably owing to the use of too few QM calculations (3600 per transformation).[13,25]

Therefore, we have here employed a system for which we can perform many more QM calculations: First, we used the octa-acid host–guest system, which is smaller and simpler than a protein. Second, we removed all eight carboxylate groups on the host molecule to further reduce the size of the system, to remove possible problems of the extensive net charge of the host, and to reduce the flexibility of the host and therefore improve the sampling of the phase space. Third, we employed the SQM PM6-DH2X method, which is computationally much cheaper than the DFT methods we have used in our previous studies. On the other hand, this means that we cannot strictly compare to any experimental results and that we use a QM method with an appreciably lower accuracy than dispersion-corrected DFT methods.[77]

We first showed that the truncation of the host gave rather restricted changes in relative binding free energies, as estimated at the MM level, up to 3 kJ/mol. Moreover, the new calculations were much more precise than those based on the full OA host, with a precision of 0.02–0.08 kJ/mol. This was partly an effect of the longer simulations (120 ns per λ value), but there were also clear indications that the sampling has been improved. In particular, our overlap measures clearly showed that the simulations were perfectly converged, in variance to the original simulations.

Next, we tested six different methods to calculate relative binding free energies at the QM level. We showed that approaches based on QM calculations for all 13 λ values (both NBB13 and QM-MBAR) were very expensive and gave poorly converged results. On the other hand, NBB4 and ssEA gave promising results, although a full convergence could not be obtained even with 1,440,000 or 720,000 QM calculations per transformation, respectively. Instead, the results indicated that 3–50 times more QM calculations are needed for convergence and this is most likely an underestimate owing to the bad conditioning of these methods (they strongly depend on one or very few of the calculated values).

However, the ssEAc approach with the cumulant expansion gave nicely converged results with a standard error of 1 kJ/mol using 720,000 QM calculations per λ transformation for all eight studied relative free energies. Moreover, it showed the expected square-root dependence of the standard error with respect to the number of calculations. It also required half as many QM calculations as the NBB4 approach. Pure averages for the MM→QM perturbation also gave converged free energies, but the results differed from those obtained by ssEAc by up to 14 kJ/mol, because this is only an approximate method that strictly should not work when the variation in the MM–QM energy differences is large.

The required number of QM calculations is of a comparable magnitude to what has been used in previous QM/MM-FES studies by Mulholland and coworkers,[24,26] especially as they included only a single rigid water molecule in the QM system. König et al. also had to perform 20,000–60,000 QM calculations to obtain an uncertainty of up to 2 kJ/mol, again including only a small solute in the QM system.[30] However, Skylaris et al. have included the solute and 200 water molecules in the QM calculations, calculating the free energies by ssEA, based on interaction energies.[27] Still, they claim to obtain converged relative solvation free energies (within 4 kJ/mol) with only 1080 QM calculations per transformation. The reason for this seems to be a smaller difference between the QM and MM potentials (although they use the same GAFF/TIP3P MM method as we do): They report a range for the E^{QM}–E^{MM} energy difference of ~55 kJ/mol, whereas it is almost four times larger in our study, 181–211 kJ/mol. The convergence of FES strongly depend on this range—in the well-converged FES calculations at the MM level (with 13 λ values), the range is typically ~10 kJ/mol with a maximum of 38 kJ/mol. The reason for the lower range in the studies by Skylaris et al. is probably that they study only simple and rigid phenol derivatives. In a recent study with more flexible (but smaller) molecules, they employed more QM calculations and included also an intermediate QM/MM step to improve the convergence with only the solute in the QM, performing QM/MM MD simulations.[31] Finally, it should also be noted that König et al. have in two studies suggested that NBB4 gives better convergence than ssEA (for absolute solvation free energies).[29,30] However, they did not employ the cumulant expansion for ssEA, which strongly improves the convergence in this study. In fact, very recently they published a study of hydration free energies of 20 organic molecules, in which they come to conclusions very similar to ours:[78] With a cumulant expansion, they obtain a slightly better convergence of the QM/MM free energies with ssEAc than with NBB4, both analytically and in practice. This shows that our conclusions apply also to other types of systems.

Consequently, we recommend the ssEAc method to obtain converged QM/MM binding free energies. Still, our results show that a very large number of QM calculations are needed to obtain strict QM/MM FES binding free energies, 720,000 per perturbation. This provides a useful guide for future studies of QM/MM binding free energies: The QM method and the size of the QM system have to be selected to allow for such an amount of QM calculations. Moreover, it provides a firm basis for comparison with alternative methods. It has recently been shown that reasonable binding free energies can be obtained with single structures optimized with dispersion-corrected DFT methods for host–guest systems;[77,79] in fact, relative energies for ligands binding to the same host were reproduced with a MAD of 5 kJ/mol. Such an approach requires only a few hundred QM energy calculations. Unfortunately, the approach worked appreciably worse for the OA system, with MADs of 5–10 kJ/mol, probably owing to the larger flexibility and the high charge of this host.[13,14] Alternatively, full QM/MM MD simulation could be performed; 720,000 QM calculations correspond to 0.36–1.4 ns simulations, depending on the time step, which may give converged FES results. Thus, it might be better to spend the QM calculations on true FES calculations with sampling at the QM/MM level instead. This is currently investigated in our group.

Acknowledgments

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.

Keywords: ligand binding · QM/MM free-energy perturbation · quantum mechanics · semi-empirical methods · octa-acid host · host–guest systems · single-step exponential averaging · non-Boltzmann Bennett acceptance ratio method

How to cite this article: M. A. Olsson, P. Söderhjelm, U. Ryde. *J. Comput. Chem.* **2016**, *37*, 1589–1600. DOI: 10.1002/jcc.24375

Additional Supporting Information are found in the online version of this article.

[1] H. Gohlke, G. Klebe, *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2644.
[2] J. Wereszczynski, J. A. McCammon, *Q. Rev. Biophys.* **2012**, *45*, 1.
[3] N. Hansen, W. F. van Gunsteren, *J. Chem. Theory Comput.* **2014**, *10*, 2632.
[4] K. Raha, M. B. Peters, B. Wang, N. Yu, A. M. Wollacott, L. M. Westerhoff, K. M. Merz, *Drug Discov. Today* **2007**, *12*, 725.
[5] P. Söderhjelm, S. Genheden, U. Ryde, In Protein-Ligand Interactions; H. Gohlke, Ed.; Wiley-VCH Verlag: Weinheim, Vol. *53*, **2012**; pp. 121–143.
[6] U. Ryde, P. Söderhjelm, *Chem. Rev.* DOI: 10.1021/acs.chemrev.5b00630.
[7] K. Raha, K. M. Merz, *J. Am. Chem. Soc.* **2004**, *126*, 1020.
[8] A. T. Cho, V. Guallar, B. J. Berne, R. Friesner, *J. Comput. Chem.* **2005**, *26*, 915.

[9] F. Gräter, S. M. Schwarzl, A. Dejaegere, S. Fischer, J. C. Smith, *J. Phys. Chem. B* **2005**, *109*, 10474.
[10] J. Fanfrlík, A. K. Bronowska, J. Rezác, O. Prenosil, J. Konvalinka, P. Hobza, *J. Phys. Chem. B* **2010**, *114*, 12666.
[11] P. Söderhjelm, J. Kongsted, U. Ryde, *J. Chem. Theory Comput.* **2010**, *6*, 1726.
[12] P. Mikulskis, S. Genheden, K. Wichmann, U. Ryde, *J. Comput. Chem.* **2012**, *33*, 1179.
[13] P. Mikulskis, D. Cioloboc, M. Andrejic, S. Khare, J. Brorsson, S. Genheden, R. A. Mata, P. Söderhjelm, U. Ryde, *J. Comput. Aided Mol. Des.* **2014**, *28*, 375.
[14] R. Sure, J. Antony, S. Grimme, *J. Phys. Chem. B* **2014**, *118*, 3431.
[15] M. Andrejic, U. Ryde, R. Mata, P. Söderhjelm, *Chem. Phys. Chem.* **2014**, *15*, 3270.
[16] M. R. Reddy, M. D. Erion, *J. Am. Chem. Soc.* **2007**, *129*, 9296.
[17] R. S. Rathore, R. N. Reddy, A. K. Kondapi, P. Reddanna, M. R. Reddy, *Theor. Chem. Acc.* **2012**, *131*, 1096.
[18] K. Swiderek, S. Marti, V. Moliner, *Phys. Chem. Chem. Phys.* **2012**, *14*, 12614.
[19] V. Luzhkov, A. Warshel, *J. Comput. Chem.* **1992**, *13*, 199.
[20] T. H. Rod, U. Ryde, *Phys. Rev. Lett.* **2005**, *94*, 138302.
[21] F. Duarte, B. A. Amrein, D. Blaha-Nelson, S. C. L. Kamerlin, *Biochim. Biophys. Acta* **2015**, *1850*, 954.
[22] G. König, S. Boresch, *J. Comput. Chem.* **2011**, *32*, 1082.
[23] F. R. Beierlein, J. Michel, J. W. Essex, *J. Phys. Chem. B* **2011**, *115*, 4911.
[24] C. J. Woods, K. E. Shaw, A. J. Mulholland, *J. Phys. Chem. B* **2015**, *119*, 997.
[25] S. Genheden, P. Söderhjelm, U. Ryde, *J. Comput. Chem.* **2015**, *36*, 2114.
[26] C. J. Woods, F. R. Manby, A. J. Mulholland, *J. Chem. Phys.* **2008**, *128*, 014109.
[27] S. J. Fox, C. Pittock, C. S. Tautermann, T. Fox, C. Christ, N. O. J. Malcolm, J. W. Essex, C. K. Skylaris, *J. Phys. Chem. B* **2013**, *117*, 9478.
[28] S. Genheden, A. I. Cabedo Martinez, M. P. Criddle, J. W. Essex, *J. Comput. Aided Mol. Des.* **2014**, *28*, 187.
[29] G. König, F. C. Pickard, Y. Mei, B. R. Brooks, *J. Comput. Aided Mol. Des* **2014**, *28*, 245.
[30] G. König, P. S. Hudson, S. Boresch, H. L. Woodcock, *J. Chem. Theory Comput.* **2014**, *10*, 1406.
[31] C. Sampson, T. Fox, C. S. Tautermann, C. Woods, C. K. A. Skylaris, *J. Phys. Chem. B* **2015**, *119*, 7030.
[32] H. Hu, W. Yang, *Annu. Rev. Phys. Chem.* **2008**, *59*, 573.
[33] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
[34] J. Heimdal, U. Ryde, *Phys. Chem. Chem. Phys.* **2012**, *14*, 12592.
[35] M. Strajbl, G. Hong, A. Warshel, *J. Phys. Chem. B* **2002**, *106*, 13333.
[36] N. V. Plotnikov, A. Warshel, *J. Phys. Chem. B* **2012**, *116*, 10342.
[37] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420.
[38] C. Cave-Ayland, C. K. Skylaris, J. W. Essex, *J. Phys. B* **2015**, *119*, 1017.
[39] K. E. Shaw, C. J. Woods, A. J. Mulholland, *J. Phys. Chem. Lett.* **2010**, *1*, 219.
[40] L. Hu, J. Eliasson, J. Heimdal, U. Ryde, *J. Phys. Chem. A* **2009**, *113*, 11793.
[41] P. Söderhjelm, U. Ryde, *J. Phys. Chem. A* **2009**, *113*, 617.
[42] L. Hu, P. Söderhjelm, U. Ryde, *J. Chem. Theory Comput.* **2013**, *9*, 640.
[43] H. S. Muddana, A. T. Fenley, D. L. Mobley, M. K. Gilson, *J. Comput. Aided Mol. Des.* **2014**, *28*, 305.
[44] C. L. D. Gibb, B. Gibb, *J. Comput. Aided Mol. Des.* **2014**, *28*, 319.
[45] J. M. Wang, R. M. Wolf, K. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
[46] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[47] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269.
[48] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902.
[49] B. H. Besler, K. M. Merz, P. A. J. Kollman, *Comput. Chem.* **1990**, *11*, 431.
[50] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, revision A. 02; Gaussian, Inc.: Wallingford, CT, **2009**.
[51] D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, T. E. Cheatham, III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York, P. A. Kollman, AMBER; University of California: San Francisco, **2015**.
[52] S. Genheden, U. Ryde, *J. Comput. Chem.* **2011**, *32*, 187.
[53] J. W. Kaus, L. T. Pierce, R. C. Walker, J. A. McCammon, *J. Chem. Theory Comput.* **2013**, *9*, 4131.
[54] T. Steinbrecher, D. L. Mobley, D. A. Case, *J. Chem. Phys.* **2007**, *127*, 214108.
[55] T. Steinbrecher, I. Joung, D. A. Case, *J. Comput. Chem.* **2011**, *32*, 3253.
[56] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[57] X. Wu, B. R. Brooks, *Phys. Lett.* **2003**, *381*, 512.
[58] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, J. R. Haak, *J Chem. Phys.* **1984**, *81*, 3684.
[59] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.
[60] B. L. Tembe, J. A. McCammon, *Comput. Chem.* **1984**, *8*, 281.
[61] C. H. Bennett, *J. Comput. Phys.* **1976**, *22*, 245.
[62] M. R. Shirts, V. S. Pande, *J. Chem. Phys.* **2005**, *122*, 144107.
[63] M. R. Shirts, J. D. Chodera, *J. Chem. Phys.* **2008**, *129*, 124105.
[64] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300.
[65] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173.
[66] M. Korth, M. Pitonák, J. Rezác, P. Hobza, *J. Chem. Theory Comput.* **2010**, *6*, 344.
[67] J. Rezác, J. Fanfrlik, D. Salahub, P. Hobza, *J. Chem. Theory Comput.* **2009**, *5*, 1749.
[68] J. Rezác, P. Hobza, *Chem. Phys. Lett.* **2011**, *506*, 266.
[69] J. J. P. Stewart, MOPAC2012, Stewart Computational Chemistry; Colorado Springs: CO.
[70] G. Hummer, A. Szabo, *J. Chem. Phys.* **1996**, *105*, 2004.
[71] J. Kästner, H. M. Senn, S. Thiel, N. Otte, W. Thiel, *J. Chem. Theory Comput.* **2006**, *2*, 452.
[72] P. Mikulskis, S. Genheden, U. Ryde, *J. Chem. Inf. Model.* **2014**, *54*, 2794.
[73] A. Bhattacharyya, *Bull. Cal. Math. Soc.* **1943**, *35*, 99.
[74] D. Wu, D. A. Kofke, *J. Chem. Phys.* **2005**, *123*, 054103.
[75] A. Pohorille, A. Jarzynski, C. Chipot, *J. Chem. Phys. B* **2010**, *114*, 10235.
[76] H. Li, W. Yang, *Chem. Phys. Lett.* **2007**, *440*, 155.
[77] R. Sure, S. Grimme, *J. Chem. Theory Comput.* **2015**, *11*, 3785.
[78] X. Jia, M. Wang, Y. Shao, G. König, B. R. Brooks, J. Z. H. Zhang, Y. Mei, *J. Chem. Theory Comput.* **2016**, *12*, 499.
[79] S. Grimme, *Chem. Eur. J.* **2012**, *18*, 9955.

Received: 2 February 2016
Revised: 8 March 2016
Accepted: 9 March 2016
Published online on 27 April 2016

Paper II

Martin A. Olsson, Ulf Ryde, *Journal of Chemical Theory and Computation,* **2017**, 13, 2245–2253

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

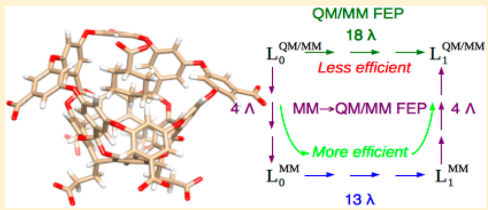
Comparison of QM/MM Methods To Obtain Ligand-Binding Free Energies

Martin A. Olsson and Ulf Ryde*

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

Supporting Information

ABSTRACT: We have compared two approaches to calculate relative binding free energies employing molecular dynamics simulations at the combined quantum-mechanical/molecular mechanics (QM/MM) level. As a test case, we study the binding of nine cyclic carboxylate ligands to the octa-acid deep-cavitand host system. The ligand is treated with the semiempirical PM6-DH+ QM method. In the first approach, we perform direct alchemical QM/MM free energy perturbation (FEP). In the second, reference-potential approach, we convert the ligands with FEP at the molecular mechanics (MM) level and then perform also MM → QM/MM FEP for each ligand. We show that the two approaches give identical results within statistical uncertainty. For the reference-potential approach, the MM → QM/MM perturbation converges in terms of energies, uncertainties, and overlap measures with two intermediate states, giving a precision of 0.5−0.9 kJ/mol for all eight transformations considered. On the other hand, the QM/MM-FEP approach requires 17−18 intermediate states, showing that the reference-potential approach is more effective. Previous calculations with single-step exponential averaging (i.e., entirely avoiding QM/MM simulations) required fewer QM/MM energy calculations, but they gave worse precision and involved approximations with an unclear effect on the results.



INTRODUCTION

One of the largest challenges of computational chemistry is to predict the binding affinities of small molecules to biological macromolecules. If such affinities could reliably be predicted with a high accuracy for any type of system, this would revolutionize drug discovery and development. Consequently, numerous methods have been suggested with this aim.[1,2] The most widely used approach is docking with simple scoring functions based on either a statistical survey of structures with known affinities or physical energy functions.[3,4] At the next step of approximation come approaches based on molecular dynamics (MD) simulations of the physical end states (the complex and possibly also the free ligand and the free macromolecule), e.g., molecular mechanics combined with Poisson−Boltzmann and surface area solvation (MM/PBSA), linear interaction energy (LIE), and similar approaches.[5−8] However, from a statistical mechanics perspective, the proper way to obtain binding free energies is by the use of free energy perturbation (FEP) methods based on MD or Monte Carlo simulations of both the end states and various intermediate states and calculating the free energies by methods such as exponential averaging (EA) (Zwanzig equation), thermodynamic integration (TI), or the (multistate) Bennett acceptance ratio (BAR) approach.[9−14] Recently, several applications of FEP to calculate relative binding free energies for large sets of protein−ligand complexes have shown a promising accuracy, at least for some proteins.[15−17] FEP methods can also be used to calculate absolute affinities by converting the ligand to a noninteracting molecule.[18−20]

FEP simulations should in principle give the correct results provided that the energy function is perfect and the sampling is infinite. Consequently, there has recently been quite some interest in improving FEP calculations of binding affinities by the use of quantum-mechanical (QM) calculations instead of the normal molecular mechanics (MM) force field.[21] Several different approaches have been employed. If only the ligand is treated with a cheap semiempirical QM method, a full FEP can be performed with the combined QM/MM potential. Only a few applications with such an approach have been presented, but often with impressive results.[22−24]

However, it has been more common to employ more accurate QM methods (e.g., density functional theory or coupled-cluster calculations) and reference-potential methods.[25−27] In these, the phase space is still sampled with MM methods, and these simulations are then used to perform an FEP in the method space from the MM to the QM/MM potential. This perturbation can in principle be performed for all intermediate states between the two ligands,[28,29] but it is more common to perform it only for the end states (i.e., the two ligands).[30,31] Then, the difference in binding energy between the two ligands is estimated by FEP at the MM level, and this energy is corrected toward the QM/MM potential by one MM → QM/MM FEP for each of the ligands. The MM → QM/MM free energy can be obtained either by reweighting the MM snapshots, e.g., by the non-Boltzmann BAR

Received: December 15, 2016
Published: March 29, 2017

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

approach (NBB)[32] or by employing single-step EA (ssEA) from MM to QM/MM.[25,28,29] The two approaches have been compared for the same system, and it has been shown that ssEA gives the more precise results both analytically and practically and actually also with a lower computational effort.[29,33,34] Alternative approaches employ energy distribution functions.[35]

The prime problem with these approaches is to obtain a proper convergence of the MM → QM/MM perturbation. For normal FEP simulations, the perturbation is divided into many small steps in order to ensure a proper overlap between the two energy distributions. However, to avoid doing the extremely time-consuming sampling at the QM/MM level, the MM → QM/MM perturbation needs to be done in a single step (employing sampling only at the MM level). If only the ligand is treated by QM or only electrostatic interactions are considered, apparently converged results have been obtained in a few cases.[30,36] However, if also the closest parts of the receptor are included in the QM region, poor precision has been observed with NBB and ssEA approaches when performing some thousands of QM calculations for each state.[28,31] Recently, we showed that 720 000 QM calculations are needed to converge the MM → QM/MM perturbation to within 1 kJ/mol for the relative binding affinities of a series of benzoate ligands to the octa-acid host−guest system.[29] Moreover, it was essential to employ a cumulant approximation[37,38] for the calculation of the ssEA free energies (i.e., assuming that the energies follow a Gaussian distribution).

Performing 720 000 QM evaluations corresponds to carrying out a fairly long MD simulation (1.4 ns with a 2 fs time step). The question therefore naturally arises whether it is better to actually perform QM/MM MD simulations to improve the convergence of the MM → QM/MM perturbation or perhaps to perform the full FEP at the QM/MM level. The advantage of such an approach is that it avoids reliance on the cumulant approximation, which is hard to confirm for this type of system. In fact, even after 720 000 QM evaluations, there were some indications that the estimate of the precision was still too optimistic, with variations of up to 9 kJ/mol in some of the free energy components among independent simulations.[29] Moreover, total energies can be employed, whereas most MM → QM/MM perturbations have used instead interaction energies, excluding differences in the internal energy of the QM region.[28,29,31,33]

QM/MM sampling has been used in a few reference-potential studies of ligand binding. Mulholland and co-workers used QM/MM Monte Carlo sampling speeded up by the Metropolis−Hastings approach[39] to study the binding of water molecules to two sites in neuraminidase (with only the water molecule treated by QM).[40] Such an approach required 160 000 QM evaluations for each state, and they observed differences of up to 12 kJ/mol between the MM and QM descriptions of the water molecules. Moreover, Skylaris and co-workers employed QM/MM simulations for the end points and one intermediate state as a stepping stone between MM and full QM calculations of the hydration energy of five small organic molecules.[36] Converged results were obtained with only 500 QM and QM/MM evaluations for each step. Boresch and co-workers showed that the MM → QM free energy can be estimated from short QM simulations (50−500 time steps but repeated 1000 times) employing nonequilibrium work functions.[41] For enzyme reactions, Warshel and co-workers have suggested the para-dynamics approach, which employs the reference-potential method but improves the convergence of the free energies by also performing the QM/MM MD simulation at the end states and calculating the free energy with the linear-response approximation.[25,42,43] However, recent studies have shown that this approximation sometimes gives poor results compared to a full FEP calculation at the QM/MM level (errors of up to 92 kJ/mol),[44] and it is not clear how it could be known beforehand whether the approximation will work (i.e., without performing the full QM/MM FEP).

In this work, we have performed calculations of relative binding free energies based on explicit QM/MM MD simulations. We investigate the convergence of reference-potential FEP calculations with up to seven intermediate states in the MM → QM/MM perturbations, discussing the accuracy of the linear-response approximation and how it can be estimated without performing additional simulations. Moreover, we compare the accuracy and efficiency of reference-potential calculations with full QM/MM FEP simulations of the same systems. As a test system, we employ the binding of nine small carboxylate ligands to the octa-acid deep cavitand,[45−47] which was employed in the SAMPL4 competition[48,49] and was also used in several of our previous studies.[28,29,50] To speed up the calculations and make it possible to compare them to full QM/MM FEP, we employed the semiempirical PM6-DH+ approach for the QM calculations,[51−53] including only the ligand in the QM system.

METHODS

Simulated System. We have studied the relative binding free energies of nine cyclic carboxylate ligands to the octa-acid deep-cavity host, which was involved in the SAMPL4 competition.[45−49] The octa-acid host and the nine ligands are shown in Figure 1. The MM calculations are taken from our previous studies of this system.[28,29] They employed the general Amber force field for both the octa-acid host and the guest molecules,[54] using symmetrized restrained electrostatic potential charges[55] obtained at the Hartree−Fock/6-31G* level of theory.[28] The TIP3P model was used to describe the water molecules.[56] As in our previous studies,[28,29] eight ligand transformations were considered: MeBz → Bz, EtBz → MeBz, pClBz → Bz, mClBz → Bz, Hx → Bz, MeHx → Hx, Hx → Pen, and Hep → Hx (employing the ligand names defined in Figure 1).

QM/MM Free Energy Simulations. Here we calculate $\Delta\Delta G^{\mathrm{bind}}_{L_0\rightarrow L_1}$, the difference between the binding free energies of pairs of ligands (L_0 and L_1) to the octa-acid host, by converting L_0 to L_1 both when the ligands are bound to the host ($\Delta G^{\mathrm{bound}}_{L_0\rightarrow L_1}$) and when they are free in solution ($\Delta G^{\mathrm{free}}_{L_0\rightarrow L_1}$):[57]

$$\Delta\Delta G^{\mathrm{bind}}_{L_0\rightarrow L_1} = \Delta G^{\mathrm{bound}}_{L_0\rightarrow L_1} - \Delta G^{\mathrm{free}}_{L_0\rightarrow L_1} \quad (1)$$

Each term on the right-hand side of this equation was estimated by FEP, slowly converting L_0 to L_1, using the energy function

$$E(\lambda) = (1-\lambda)E_0 + \lambda E_1 \quad (2)$$

where E_0 and E_1 are the potentials of the two ligands and λ is a coupling parameter that goes from 0 to 1. Our aim is to calculate each term $\Delta G^{\mathrm{s}}_{L_0\rightarrow L_1}$ (where s = bound or free) at the QM/MM level. We will compare two approaches, as illustrated in Figure 2. In the first, all of the MD simulations at different λ values were performed at the QM/MM level. This is illustrated by the upper horizontal part of the thermodynamic cycle in Figure 2. We will call this approach QM/MM-FEP.

In the second approach, we tried to decrease the number of demanding QM/MM MD simulations by employing the full

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

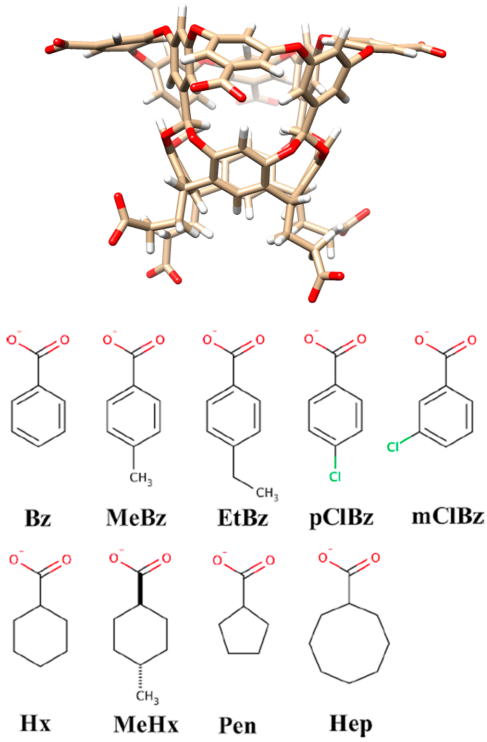


Figure 1. Octa-acid host and the nine guest molecules considered in this study.

thermodynamic cycle in Figure 2. Thus, the MD simulations at different λ values were performed at the MM level, giving $\Delta\Delta G^{s,MM}_{L_0\rightarrow L_1}$ (lower horizontal path in Figure 2). These values were calculated by employing 13 λ values, viz., 0.0, 0.05, 0.1, 0.2, ..., 0.8, 0.9, 0.95, and 1.0. Then at the end states (L_0 and L_1) we performed additional FEP calculations in the method space, going from an MM description to a QM/MM description ($\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_i}$ for i = 0, 1). On the basis of the thermodynamic cycle in Figure 2, the net QM/MM free energy can be estimated from these three paths:

$$\Delta\Delta G^{s,QM/MM}_{L_0\rightarrow L_1} = \Delta\Delta G^{s,MM}_{L_0\rightarrow L_1} - \Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_0} + \Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_1} \quad (3)$$

In previous studies, we attempted to avoid any QM/MM MD simulations by calculating the free energy differences $\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_i}$ (i = 0, 1) on the basis of only a single MD simulation at the MM level using the ssEA approach.[28,29,31] However, as discussed in the Introduction, this led to severe convergence problems, requiring 720 000 QM calculations for proper convergence.[29] Therefore, we here study whether the convergence of the MM → QM/MM perturbation can be improved by performing simulations with intermediate states. In analogy with eq 2, we form the energy function

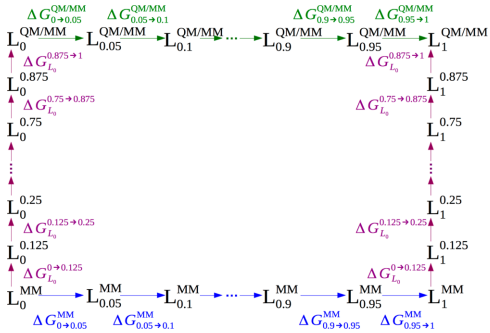


Figure 2. Thermodynamic cycle employed to study the relative free energy for the binding of two ligands (L_0 and L_1) to a target receptor. The cycle applies both when the ligand is bound to the receptor and when it is free in solution. We aim to estimate the free energy difference for converting L_0 to L_1 at the QM/MM level, either by directly calculating it by QM/MM MD simulations at a number of different λ values (top horizontal reaction with green arrows) or by converting L_0 to L_1 at the MM level (bottom horizontal reaction, blue arrows) combined with MM → QM/MM perturbations for each of the two ligands (left and right vertical reactions, magenta arrows), again by employing a number of intermediate Λ states.

$$E(\Lambda) = (1-\Lambda)E_{MM} + \Lambda E_{QM/MM} \quad (4)$$

where E_{MM} is the MM energy, $E_{QM/MM}$ is the QM/MM energy, and Λ is the coupling parameter (we use the symbol Λ here to distinguish it from the λ parameter used to convert L_0 to L_1 in eq 2). In previous studies, we performed MD simulations only for Λ = 0, whereas in this study we use two, three, five, or nine Λ values at regular intervals (i.e., Λ = 0.0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, and 1.0 for nine Λ values and a subset of these values for the other three sets). In addition, we also tried four Λ values at Λ = 0.0, 0.333, 0.666, and 1.0.

Details of the Calculations. All of the FEP and MD simulations were performed with the sander module of the Amber 14 software.[58] The QM/MM calculations employed the QM/MM implementation in this software[58,59] with electrostatic embedding. For the explicit QM/MM FEP simulations (top reaction in Figure 2), we used a dual-topology scheme with the energy function

$$E^{QM/MM}(\lambda) = (1-\lambda)E_0^{QM/MM} + \lambda E_1^{QM/MM} \quad (5)$$

Thus, for each λ value, two separate QM calculations were performed, one for L_0 and one for L_1. Then the energies were combined using eq 5, which shows how energies for the mixed states ($0 < \lambda < 1$) are naturally obtained.

To speed up the calculations and make full QM/MM FEP calculations possible, only the guest molecule was included in the QM region. The QM calculations were performed at the semiempirical PM6-DH+ level of theory.[51−53] The QM/MM FEP simulations were run on 20 CPU cores at a speed of 1−5 ns/day.

Free energy differences were calculated with the multistate Bennett acceptance ratio (MBAR) approach using the pyMBAR software.[60] This software also gives uncertainties of the calculated free energies and selects a noncorrelated subset of data points. For comparison and to evaluate the convergence, the

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

Table 1. $\Delta\Delta G_L^{MM\rightarrow QM/MM}$ QM/MM Free Energy Corrections (in kJ/mol) for the Eight Transformations Calculated with Two, Three, Four, Five, or Nine Λ Values for Both L_0 and L_1; The Last Row Shows the Mean Absolute Deviation (MAD) from the Calculations with Nine Λ Values

	ligand	two Λ values	three Λ values	four Λ values	five Λ values	nine Λ values
MeBz → Bz	MeBz	1.9 ± 4.6	−0.9 ± 1.0	−0.6 ± 0.4	−0.9 ± 0.7	−1.3 ± 0.5
	Bz	−1.5 ± 5.0	0.0 ± 0.9	−0.3 ± 0.4	−0.4 ± 0.6	−1.0 ± 0.5
EtBz → MeBz	EtBz	1.3 ± 4.9	2.1 ± 1.0	1.9 ± 0.4	2.0 ± 0.7	2.1 ± 0.5
	MeBz	3.0 ± 3.3	1.1 ± 1.0	−0.3 ± 0.4	−0.8 ± 0.6	−0.8 ± 0.5
pClBz→Bz	pClBz	−3.4 ± 2.6	−5.2 ± 0.9	−5.0 ± 0.3	−4.8 ± 0.6	−5.5 ± 0.5
	Bz	−4.5 ± 7.2	−1.4 ± 1.0	−0.3 ± 0.4	−1.1 ± 0.6	−0.9 ± 0.5
mClBz → Bz	mClBz	−13.5 ± 4.7	−4.2 ± 1.0	−4.1 ± 0.4	−3.6 ± 0.7	−3.8 ± 0.5
	Bz	−3.3 ± 3.4	−1.4 ± 1.0	−0.8 ± 0.4	−0.6 ± 0.6	−1.0 ± 0.5
Hx → Bz	Hx	1.4 ± 2.4	2.3 ± 0.9	2.9 ± 0.3	2.8 ± 0.6	2.9 ± 0.4
	Bz	−0.6 ± 2.3	−0.9 ± 0.9	−0.7 ± 0.4	−1.4 ± 0.6	−0.8 ± 0.5
MeHx → Hx	MeHx	−2.9 ± 2.2	−0.8 ± 0.9	−0.6 ± 0.3	−1.1 ± 0.6	−0.6 ± 0.4
	Hx	3.7 ± 2.5	2.3 ± 0.8	2.1 ± 0.4	3.0 ± 0.6	2.7 ± 0.4
Hx→Pen	Pen	4.0 ± 3.2	3.4 ± 0.9	2.9 ± 0.3	3.1 ± 0.6	2.4 ± 0.5
	Hx	0.7 ± 2.3	2.3 ± 0.9	3.3 ± 0.3	2.0 ± 0.6	2.7 ± 0.4
Hep → Hx	Hep	−4.0 ± 3.0	1.8 ± 0.9	2.1 ± 0.4	0.6 ± 0.6	0.9 ± 0.5
	Hx	7.0 ± 2.4	3.5 ± 0.9	3.6 ± 0.3	3.0 ± 0.6	3.0 ± 0.4
MAD		2.7	0.5	0.4	0.4	

free energies were calculated also with BAR, TI, and EA using the same software.[10−12]

The octa-acid host and the guest molecules were solvated in a truncated octahedral box of explicit water molecules extending at least 10 Å from the solute using the leap program in the Amber suite, giving ∼5000 and ∼1700 atoms in total for the calculations with and without the host, respectively. No counterions were employed—our previous test calculations indicated that they have only a small effect on the relative binding free energies (<2 kJ/mol).[28]

We first performed 500 cycles of minimization with the heavy atoms of the host and guest molecules restrained toward the starting structure with a force constant of 418.4 kJ mol^{-1} $Å^{-2}$. This was followed by 20 ps of constant-volume equilibration with the same restraints and 0.5 ns of constant-pressure equilibration without any restraints. Finally, a 1 ns production simulation was run, during which energies were sampled every 5 ps (every 1 ps for the simulations with four Λ values).

The temperature was kept constant at 300 K using Langevin dynamics with a collision frequency of 2 ps^{-1}.[61] The pressure was kept constant at 1 atm using Berendsen's weak-coupling isotropic algorithm with a relaxation time of 1 ps.[62] Long-range electrostatics were handled by particle-mesh Ewald summation[63] with a fourth-order B-spline interpolation and a tolerance of 10^{-5}. The cutoff radius for Lennard−Jones interactions was set to 8 Å. The MM calculations were performed with all bonds involving hydrogen atoms constrained to their equilibrium values using the SHAKE algorithm,[64] allowing for a time step of 2 fs during the simulations. However, this was not allowed within the QM region. Therefore, all of the QM/MM MD simulations were run without any constrained bonds and a time step of 1 fs. Test QM/MM-FEP simulations showed that a time step of 0.5 fs gave identical results (within 0.1 kJ/mol for the pClBz → Bz transformation).

Quality Measures. To study the convergence for the various perturbations, several overlap measures were employed, viz., the Bhattacharyya coefficient (Ω),[65] the Wu and Kofke overlap measures of the energy probability distributions (K_{AB}),[66] and their bias metrics (Π).[66] The quality of the FEP calculations was further analyzed using the weight of the maximum term in the exponential average (w_{max}),[27] the difference of the forward and backward exponential average estimates ($\Delta\Delta G_{EA}$), and the difference between the BAR and TI estimates ($\Delta\Delta G_{TI}$).[16]

The performance of the free energy estimates compared with experimental data was quantified by the mean absolute deviation (MAD), the correlation coefficient (R^2), and Kendall's rank correlation coefficient (τ_r). The last of these was calculated only for the transformations that were explicitly studied, not for all combinations that could be formed from these transformations. The standard deviations of the quality measures were obtained by a simple simulation approach:[67] Each transformation was assigned a random number from a Gaussian distribution with the mean and standard deviation of the mean obtained from the BAR calculations. The quality measures where then calculated, and the procedure was repeated 1000 times.

RESULTS AND DISCUSSION

In this work, we have compared two methods to estimate the difference in binding free energy of two ligands to a receptor at the QM/MM level of theory using FEP calculations. As a test system, we have used the binding of nine cyclic carboxylate ligands to the octa-acid deep-cavity host.[40] The experimental data were taken from the SAMPL4 challenge.[48,49] The octa-acid host is a macrocycle of 184 atoms with fourfold symmetry, 12 benzene rings, and eight carboxylic groups (Figure 1). It forms a hydrophobic cavity with a depth of 10 Å that can bind various small molecules in aqueous solution. The nine ligands are also shown in Figure 1. They are all cyclic (with five to seven atoms in the ring) with a carboxylate group and therefore a single negative charge. We studied eight ligand transformations: MeBz → Bz, EtBz → MeBz, pClBz → Bz, mClBz → Bz, Hx → Bz, MeHx → Hx, Hx → Pen, and Hep → Hx. These mostly involve H → methyl or H → Cl substitutions, although two involve the deletion of a CH_2 group inside the ring (Hx → Pen and Hep → Hx) and one involves the conversion of cyclohexane to benzene (Hx → Bz). We have performed two types of FEP calculations, which are described in separate sections.

Reference-Potential Calculations. We first tried to calculate the relative binding free energies by the reference-potential approach using the full thermodynamic cycle in Figure

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

2. Thus, we first calculated $\Delta\Delta G^{MM}_{L_0\rightarrow L_1}$ at the MM level (lower horizontal branch in Figure 2). This involved normal FEP calculations at the MM level, and these results were taken from our previous MM study of the same system and will not be much discussed.[28] Then we calculated the free energies for conversion of each of the two ligands from an MM description to a QM/MM description, both when the ligands are bound to the host or when they are free in solution ($\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_i}$ for $L_i = L_0$ or L_1 and s = bound or free; i.e., the two vertical branches in Figure 2).

In previous studies, we and others have tried to estimate this term with ssEA or NBB, thereby avoiding MD simulations at the QM/MM level.[28−31,33] However, it has been observed that the convergence is very poor, requiring at least 720 000 QM calculations.[29] Here we investigate whether the convergence can be improved by performing QM/MM MD simulations at the end points and possibly at additional intermediate states, described by the Λ coupling parameter in eq 4. We have tested calculations with two, three, four, five, and nine Λ values.

The individual $\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L}$ terms were rather large, ranging from −478 to −768 kJ/mol (Table S1). However, for each ligand, the values were of a similar size when the ligand was bound to the host or free in water solution. Therefore, the difference between these two calculations, $\Delta\Delta G^{MM\rightarrow QM/MM}_{L} = \Delta G^{bound,MM\rightarrow QM/MM}_{L} - \Delta G^{free,MM\rightarrow QM/MM}_{L}$, which is shown in Table 1, was much smaller (e.g., −5.5 to 3 kJ/mol with nine Λ values).

From the results in Table 1, it can directly be seen that the calculations with two Λ values gave unreliable results. The calculated $\Delta\Delta G^{MM\rightarrow QM/MM}_{L}$ values differed by up to 10 kJ/mol from the results obtained with nine Λ values (for the individual $\Delta G^{s,MM\rightarrow QM/MM}_{L}$ results in Table S1, the difference was somewhat smaller, up to 6 kJ/mol). The calculations with three, four, and five Λ values gave appreciably smaller differences, up to 1.8, 1.5, and 0.8 kJ/mol, respectively, with average differences of 0.5, 0.4, and 0.4 kJ/mol, respectively.

Likewise, the standard errors of the calculated $\Delta\Delta G^{MM\rightarrow QM/MM}_{L}$ free energies were also large and varying with two Λ values (2−7 kJ/mol). With more Λ values, the precision improved strongly: already with three Λ values it was 0.8−1.0 kJ/mol for all of the calculations, and it decreased to 0.4−0.5 kJ/mol for nine Λ values. The precision with four Λ values was even better (0.3 kJ/mol) because energy data were collected five times more often (every 1 ps instead of every 5 ps). The free energies were obtained from a noncorrelated subset of the calculated energies, estimated using the pyMBAR software.[60] However, omission of correlated data changed the resulting free energies by less than 0.8 kJ/mol, indicating that the sampling frequency of 5 ps in the other calculations was unnecessarily sparse (the correlation time of the potential energy differences estimated by pyMBAR was 1−6 ps, with averages of 1.6 and 1.2 ps for the bound and free simulations, respectively).

These results were obtained with the MBAR method. Quite satisfyingly, BAR calculations gave essentially identical results, with an average deviation of only 0.04 kJ/mol (maximum difference 0.2 kJ/mol). However, TI and EA calculations gave larger deviations of 0.7 and 2.9 kJ/mol on average (maximum 8 and 24 kJ/mol), respectively. The difference was largest for the calculations with two Λ values, 2.7 and 6.7 kJ/mol on average, and decreased to 0.1 and 0.3 kJ/mol for nine Λ values. This reflects the poorer overlap with fewer Λ values, which is partly amended by the use of BAR or MBAR.

Further estimates of the convergence can be obtained by observing the fact that the $\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_i}$ free energies depend only on a single ligand (i.e., they do not depend on the other ligand in the perturbation). This means that we actually have four independent estimates of $\Delta\Delta G^{s,MM\rightarrow QM/MM}_{L_i}$ for the Bz and Hx ligands and two estimates for the MeBz ligand in Table S1. For the individual calculations with two Λ values, the results varied by up to 6.6 kJ/mol with a standard deviation of 2.0 kJ/mol. With three, four, five, and nine Λ values, the standard deviation decreased to 0.6, 0.5, 0.4, and 0.2 kJ/mol, respectively, and the maximum deviation was 2.6, 2.5, 1.3, and 0.8 kJ/mol, respectively. The results were similar when taking the difference of the calculations with the ligand bound to the host or free in solution, as shown in Table 1 (e.g., the standard deviation was 1.9, 0.7, 0.7, 0.4, and 0.1 kJ/mol with two, three, four, five, and nine Λ values, respectively).

Thus, the results in Tables 1 and S1 show that two Λ values are too few to obtain reliable results for the MM → QM/MM perturbation; instead, three to five Λ values should be used, depending on the desired accuracy and precision. However, it is likely that this conclusion may depend on what system is studied. Therefore, it would be strongly desirable if we could decide directly from the calculations with a certain number of Λ values whether the results are reliable. We have seen that the standard error of the calculated free energies give a first indication in this direction.

To obtain more quantitative measures, we tested six overlap measures that have been employed to judge the overlap in previous FEP calculations,[16,29] viz., the Bhattacharyya coefficient (Ω),[65] the Wu and Kofke overlap measures (K_{AB})[66] and their bias metrics (Π),[66] the weight of the maximum term in the exponential average (w_{max}),[27] the difference of the forward and backward exponential average estimates ($\Delta\Delta G_{EA}$), and the difference between the BAR and TI estimates ($\Delta\Delta G_{TI}$).[16] To start with, we used the following indicators of poor overlap: Ω < 0.7, K_{AB} < 0.7, Π < 0.5, w_{max} > 0.3, $\Delta\Delta G_{EA}$ > 4 kJ/mol, and $\Delta\Delta G_{EA}$ > 4 kJ/mol.[16,66] The results were quite clear and consistent: For the calculations with two and three Λ values, several measures indicated poor overlap, in particular Π and w_{max} and in most cases also $\Delta\Delta G_{EA}$. On the other hand, most of the criteria indicated proper overlap for the calculations with four, five, and nine Λ values. The only exception was w_{max}, which sometimes indicated a poor overlap for the calculations with four or five Λ values.

Except for the Π measure,[66] the overlap criteria used are rather arbitrary, without any theoretical basis. Considering that the results were so consistent, we decided to determine improved limits for the various overlap measures based on these results, optimizing the consistency with the results of the Π measure, i.e., that all calculations with two and three Λ values show poor overlap whereas those with four, five, and nine Λ values have acceptable overlap. For Ω, the best results were obtained between 0.84 and 0.85, for which only three of the 64 calculations gave results that were not consistent with the Π measure. K_{AB} attains values between 0 and 2, and in the present study they are always larger than unity. A double-sided criterion seems to be necessary. In the present study, values between 1.25 and 1.27 gave only three calculations that disagreed with the Π measure. Thus, $|K_{AB} - 1| > 0.25$ seemed to be a proper criterion, also agreeing with Ω. For w_{max}, the original criterion was apparently too low. Instead, the best agreement with the Π measure was obtained between 0.47 and 0.60, for which only three

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

Table 2. Overlap Measures for MM → QM/MM Perturbations of the Various Transformations for Various Numbers of Λ Values (2−9)[a]

		L_0						L_1					
	no. of Λ values	Ω	K_{AB}	Π	w_{max}	$\Delta\Delta G_{EA}$	$\Delta\Delta G_{TI}$	Ω	K_{AB}	Π	w_{max}	$\Delta\Delta G_{EA}$	$\Delta\Delta G_{TI}$
MeBz → Bz	2	**0.4**	**1.7**	**−3.0**	**1.00**	**14.3**	0.1	**0.4**	**1.7**	**−2.3**	**0.99**	**35.0**	**5.0**
	3	**0.7**	1.3	**−0.8**	**0.85**	**3.7**	0.2	**0.7**	**1.3**	**−0.5**	**0.68**	**5.9**	**0.3**
	4	0.9	1.3	0.7	0.23	0.8	0.0	0.9	1.3	0.8	0.32	0.4	0.0
	5	0.9	1.1	0.5	0.29	0.8	0.1	0.9	1.1	0.9	0.20	1.1	0.2
	9	0.9	1.1	1.4	0.06	0.4	0.0	0.9	1.0	1.4	0.06	0.5	0.1
EtBz → MeBz	2	**0.7**	**1.7**	**−2.1**	**0.93**	**35.0**	**3.2**	**0.6**	**1.7**	**−2.4**	**1.00**	**22.8**	**1.5**
	3	0.9	**1.4**	**−0.3**	**0.62**	**3.8**	**0.9**	0.9	**1.3**	**−0.4**	**0.82**	**4.5**	**0.3**
	4	0.9	1.3	0.5	0.68	3.9	0.0	1.0	1.2	0.8	0.24	0.9	0.0
	5	0.9	1.1	0.9	0.17	1.0	0.3	0.9	1.1	0.9	0.23	1.3	0.1
	9	0.9	1.0	1.6	0.04	0.4	0.1	0.9	1.1	1.3	0.16	0.5	0.0
pClBz → Bz	2	**0.7**	**1.7**	**−2.3**	**0.87**	**23.4**	0.1	**0.6**	**1.6**	**−2.4**	**1.00**	**39.9**	**4.1**
	3	**0.8**	**1.3**	**−0.5**	**0.72**	**3.2**	**0.6**	**0.8**	**1.3**	**−0.4**	**0.81**	**7.2**	**0.3**
	4	1.0	1.3	0.8	0.17	1.1	0.0	1.0	1.2	1.0	0.12	1.5	0.0
	5	0.9	1.1	0.7	0.23	1.1	0.1	0.9	1.1	0.9	0.21	1.4	0.2
	9	0.9	1.0	1.4	0.08	0.7	0.0	0.9	1.1	1.6	0.08	0.5	0.0
mClBz → Bz	2	**0.6**	**1.7**	**−2.3**	**1.00**	**34.8**	**8.3**	**0.7**	**1.6**	**−2.3**	**0.94**	**26.9**	**3.0**
	3	**0.8**	**1.3**	**−0.7**	**0.92**	**5.0**	**0.5**	**0.8**	**1.3**	**−0.7**	**0.93**	**7.0**	**0.6**
	4	1.0	1.2	0.8	0.25	1.5	0.0	1.0	1.2	0.8	0.15	1.0	0.0
	5	0.9	1.1	0.8	0.35	1.8	0.1	0.9	1.1	0.8	0.49	1.5	0.0
	9	0.9	1.1	1.4	0.15	0.5	0.1	0.9	1.0	1.6	0.08	0.4	0.0
Hx → Bz	2	**0.5**	**1.6**	**−2.3**	**0.99**	**18.7**	**1.9**	**0.4**	**1.6**	**−2.9**	**1.00**	**13.0**	**0.7**
	3	**0.8**	**1.3**	**−0.6**	**0.87**	**3.3**	0.1	**0.7**	**1.3**	**−0.8**	**0.91**	**3.4**	**0.8**
	4	0.9	1.2	0.8	0.22	1.1	0.0	0.9	1.2	0.6	0.22	0.9	0.0
	5	0.9	1.1	0.6	0.33	1.4	0.0	0.9	1.1	0.6	0.38	1.0	0.1
	9	0.9	1.0	1.5	0.06	0.5	0.0	0.9	1.0	1.5	0.08	0.6	0.1
MeHx → Hx	2	**0.7**	**1.6**	**−2.0**	**0.74**	**17.3**	**3.3**	**0.6**	**1.6**	**−1.8**	**0.83**	**19.8**	**1.9**
	3	0.9	**1.3**	**−0.3**	**0.81**	**2.2**	**0.4**	**0.8**	1.2	**−0.3**	0.43	**4.3**	**0.7**
	4	1.0	1.2	0.8	0.24	1.0	0.0	0.9	1.1	0.9	0.29	2.6	0.0
	5	0.9	1.1	0.9	0.22	0.6	0.2	0.9	1.1	0.9	0.22	0.8	0.0
	9	0.9	1.0	1.5	0.07	0.3	0.1	0.9	1.0	1.6	0.05	0.3	0.0
Hx → Pen	2	**0.5**	**1.6**	**−2.2**	**0.95**	**28.1**	**1.2**	**0.4**	**1.7**	**−2.1**	**0.95**	**16.7**	**1.0**
	3	**0.8**	**1.3**	**−0.2**	**0.61**	**4.8**	**0.5**	**0.8**	**1.3**	**−0.3**	**0.70**	**2.5**	0.1
	4	0.9	1.1	0.8	0.25	1.1	0.0	0.9	1.2	0.9	0.14	0.9	0.0
	5	0.9	1.1	0.9	0.22	0.6	0.1	0.9	1.1	0.8	0.17	0.9	0.2
	9	0.9	1.0	1.5	0.09	0.4	0.1	0.9	1.0	1.6	0.06	0.3	0.0
Hep → Hx	2	**0.5**	**1.6**	**−2.1**	**0.97**	**24.8**	**4.2**	**0.5**	**1.6**	**−1.9**	**0.82**	**18.7**	**2.9**
	3	**0.8**	**1.3**	**−0.2**	**0.65**	**6.4**	**0.3**	**0.8**	1.3	−0.3	0.37	**4.3**	**0.4**
	4	0.9	1.2	1.0	0.12	1.0	0.0	0.9	1.1	0.9	0.23	0.6	0.0
	5	0.9	1.1	0.9	0.46	2.2	0.1	0.9	1.1	0.8	0.13	1.1	0.2
	9	0.9	1.0	1.4	0.08	0.5	0.0	0.9	1.0	1.6	0.12	0.6	0.1

[a]The overlap measures are described in the main text. Values that violate the new criteria (Ω > 0.85, $|K_{AB} - 1| < 0.25$, Π > 0.5, $w_{max} < 0.5$, $\Delta\Delta G_{EA} < 2$ kJ/mol, and $\Delta\Delta G_{TI} < 0.25$ kJ/mol) are marked in bold type.

calculations gave a diverging conclusion. For $\Delta\Delta G_{EA}$, all of the values between 1.8 and 3.2 kJ/mol gave three discrepancies compared with the Π measure. $\Delta\Delta G_{TI}$ seemed to be the most problematic and least useful criterion. Clearly, 4 kJ/mol was much too large, but all of the values between 0.21 and 0.29 kJ/mol gave six discrepancies compared with the Π measure. This is probably the case because these low levels are close to the precision of the estimates and they are also affected by integration errors in TI. Therefore, we do not recommend $\Delta\Delta G_{TI}$ as an overlap criterion. For the other measures, we suggest the following improved criteria for proper overlap: Ω > 0.85, $|K_{AB} - 1| < 0.25$, Π > 0.5, $w_{max} < 0.5$, and $\Delta\Delta G_{EA} < 2$ kJ/mol. Values violating these criteria are shown in bold face in Table 2. Future studies on other systems are needed to see how general these criteria are.

Finally, we combined the $\Delta\Delta G_{L_i}^{MM\rightarrow QM/MM}$ estimates with the MM estimates of the binding free energies ($\Delta\Delta G_{L_0\rightarrow L_1}^{MM}$ in eq 3)[28] to get our final QM/MM estimates of the binding affinities. These results are collected in Table 3. The MM → QM corrections for the relative binding affinities (with nine Λ values; $\Delta\Delta G_{L_1}^{MM\rightarrow QM/MM} - \Delta\Delta G_{L_0}^{MM\rightarrow QM/MM}$) were rather small (−4 to +5 kJ/mol), but the estimated relative affinities are also rather small (2−13 kJ/mol), so the effect of the correction is quite significant.

As for the MM → QM/MM corrections, the final $\Delta\Delta G_{bind}$ estimates obtained with two Λ values were highly uncertain and inaccurate, with standard errors of 3−8 kJ/mol. With more Λ values, the precision improved from 1.2−1.4 kJ/mol (three Λ) to 0.7−1.0 kJ/mol (nine Λ). The results with three, four, and five Λ

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

Table 3. Final QM/MM Estimates of the Relative Binding Affinities (in kJ/mol) Obtained by Reference-Potential Calculations with Two to Nine Λ Values and by Direct QM/MM-FEP Calculations (Column "horiz."); For Comparison, Experimental Affinities[49] and Our Results Obtained at the MM Level[28] Are Also Included, Along with MAD, R^2, and τ_r Measures of the Various Sets of Estimates Compared with the Experimental Affinities

	MM	two Λ values	three Λ values	four Λ values	five Λ values	nine Λ values	horiz.	exptl
MeBz → Bz	15.9 ± 0.1	12.5 ± 6.8	16.8 ± 1.4	16.3 ± 0.5	16.4 ± 0.9	16.3 ± 0.7		8.9 ± 0.5
EtBz → MeBz	1.0 ± 0.1	2.7 ± 5.9	0.0 ± 1.4	−1.2 ± 0.5	−1.7 ± 0.9	−1.9 ± 0.7		1.7 ± 0.5
pClBz → Bz	19.1 ± 0.1	18.0 ± 7.6	22.8 ± 1.3	23.8 ± 0.5	22.7 ± 0.9	23.7 ± 0.7	23.2 ± 0.9	12.5 ± 0.2
mClBz → Bz	7.1 ± 0.1	17.4 ± 5.8	9.9 ± 1.4	10.3 ± 0.6	10.2 ± 0.9	9.9 ± 0.7	11.3 ± 0.8	6.4 ± 0.3
Hx → Bz	13.1 ± 0.3	11.1 ± 3.4	9.9 ± 1.3	9.5 ± 0.6	9.0 ± 0.9	9.4 ± 0.7		7.9 ± 0.4
MeHx → Hx	15.4 ± 0.2	21.9 ± 3.4	18.5 ± 1.2	18.1 ± 0.6	19.4 ± 0.9	18.7 ± 0.6		8.3 ± 0.4
Hx → Pen	7.5 ± 0.7	4.1 ± 4.0	6.5 ± 1.4	7.9 ± 0.9	6.3 ± 1.1	7.8 ± 1.0		7.9 ± 0.4
Hep → Hx	5.5 ± 0.7	16.5 ± 3.9	7.1 ± 1.4	7.0 ± 0.8	7.9 ± 1.1	7.5 ± 0.9		4.1 ± 0.3
MAD	3.6 ± 0.2	6.6 ± 1.5	4.9 ± 0.5	4.7 ± 0.2	5.2 ± 0.4	5.0 ± 0.3		
R^2	0.84 ± 0.04	0.21 ± 0.16	0.79 ± 0.07	0.86 ± 0.04	0.76 ± 0.06	0.83 ± 0.05		
τ_r	1.00 ± 0.00	1.00 ± 0.15	0.75 ± 0.13	0.75 ± 0.01	0.75 ± 0.04	0.75 ± 0.01		

values agreed with those obtained with nine Λ values within 1.9, 0.5, and 0.8 kJ/mol (0.7, 0.5, and 0.6 kJ/mol on average), respectively.

The calculated relative affinities can be compared with experimental affinities,[49] which are also included in Table 3. The results were quite good, giving (with nine Λ values) a mean absolute deviation (MAD) of 5.0 ± 0.3 kJ/mol, a correlation coefficient (R^2) of 0.83 ± 0.5, and the correct sign for all of the affinities except the smallest one (τ_r = 0.75 ± 0.01). The results with only two Λ values were much worse, with R^2 = 0.21 ± 0.16 and MAD = 6.6 ± 1.5 kJ/mol. However, the results with three to nine Λ values were statistically equivalent (R^2 = 0.76−0.86 and MAD = 4.7−5.2 kJ/mol).

The present results are somewhat better than our previous QM/MM FEP results obtained with ssEA, which gave no correlation (R^2 = 0.00 ± 0.03) but a similar MAD of 4.9 ± 0.4 kJ/mol.[29] However, the results are still not better than the original MM results:[28] the correlation coefficient is similar (R^2 = 0.83 compared to 0.84), but the MAD is slightly worse (5.0 ± 0.3 kJ/mol compared with 3.6 ± 0.2 kJ/mol). In fact, as can be seen in Figure 3, the QM/MM calculations improved the free energies of only two of the eight transformations.

Direct QM/MM FEP. For comparison, we also performed some full QM/MM FEP calculations, i.e., by performing all of the MD simulations at the QM/MM level and calculating the QM/MM free energy difference $\Delta\Delta G_{bind}$ directly using the upper horizontal route in Figure 2 (i.e., using no MM simulations). For technical reasons, these calculations could only be performed for the two Cl → H perturbations (i.e., the pClBz → Bz and mClBz → Bz perturbations). The overlap measures indicated problems for many of the calculations with large λ values (λ = 0.6−1.0). This was solved by adding simulations at four or five additional intermediate λ values (0.65, 0.75, 0.85, 0.93, and 0.98).

The results of these calculations are also included in Table 3. It can be seen that the results agree reasonably well with the calculations with the reference potential: For the pClBz → Bz perturbation, $\Delta\Delta G_{bind}$ = 23.2 ± 0.9 kJ/mol, compared to 23.7 ± 0.7 kJ/mol with the reference potential and nine Λ values. For the mClBz → Bz perturbation, the two results were 11.3 ± 0.8 and 9.9 ± 0.7 kJ/mol. This indicates that the two sets of results were consistent and that the calculations were converged.

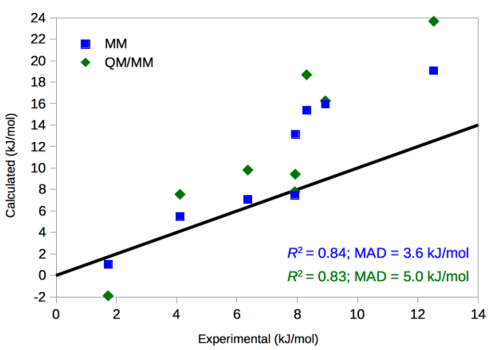


Figure 3. Comparison of the experimental and calculated affinities obtained with either MM or the reference-potential QM/MM method with nine Λ values. The line shows the perfect correlation.

CONCLUSIONS

In this paper, we have compared two approaches to obtain relative binding free energies at the QM/MM level, viz., by direct QM/MM FEP calculations (upper horizontal route in Figure 2) and with the reference-potential approach (the other three routes in Figure 2). In variance to most previous approaches,[28−31,33] we have improved the convergence of the MM → QM/MM perturbations by actually performing QM/MM MD simulations for a number Λ values. We have shown that the two approaches give consistent results within the statistical uncertainty, provided that the convergence of both approaches is carefully monitored and that proper overlap between states is ensured.

In particular, we have studied how many Λ values are needed for the reference-potential calculations. The results in Tables 1 and S1 clearly show that two Λ values are not enough to obtain any reliable results, giving errors of 2−9 kJ/mol in the final free energies and a statistical precision of 3−8 kJ/mol. TI and EA calculations gave even larger errors with few Λ values. This shows that a linear-response approximation (i.e., the TI results with two Λ values) is inappropriate for the present systems. This approximation is often used in the paradynamics approach without any test of the overlap.[26,42−44] In fact, our results strongly suggest that MBAR or BAR calculations should be employed for this type of calculation, where the overlap is often problematic.

Already with three Λ values the results are much improved, showing errors of less than 2 kJ/mol (0.7 kJ/mol on average)

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

compared to those with nine Λ values and a precision of 1.2−1.4 kJ/mol. However, a number of overlap entries indicate that it may be dangerous to trust these results (cf. Table 2). Instead, we tend to prefer the results obtained with four Λ values, for which the overlap is proper, the maximum and average errors are reduced to 0.6 and 0.3 kJ/mol, and the precision is 0.5−0.9 kJ/mol (with a sampling frequency of 1 ps).

We strongly emphasize the importance of monitoring the overlap for the MM → QM/MM perturbations. We have tested six different overlap criteria: Ω, K_{AB}, Π, w_{max}, $\Delta\Delta G_{EA}$, and $\Delta\Delta G_{TI}$. In accordance with our previous studies,[16] Π seemed to be the most reliable measure, whereas $\Delta\Delta G_{TI}$ cannot be recommended. The results were so consistent that they could be used to suggest a new set of stricter criteria for proper overlap: Ω > 0.85, $|K_{AB} - 1| < 0.25$, Π > 0.5, $w_{max} < 0.5$, and $\Delta\Delta G_{EA} < 2$ kJ/mol. Moreover, the MBAR precision also provides a useful guide whether the calculations are reliable. It is possible that these values may depend on the simulated system, and the same most likely also applies to the number of Λ values needed to obtain converged results. Therefore, great caution is recommended when studying a new system, carefully following the convergence of the free energies with the help of the overlap measures and comparing the results obtained with different numbers of Λ values.

Most importantly, the present calculations allow us to compare the efficiencies of various approaches to calculate binding free energies at the QM/MM level. The reference-potential approach with four Λ values required in total 12 QM/MM MD simulations (three simulations each for the two ligands, both when bound to the host and when free in solution; the Λ = 0 simulation can be performed at the MM level). The direct QM/MM FEP instead required 34−36 λ values because of the poor overlap observed at high λ values. This indicates that the reference-potential approach actually is ∼3 times more effective than the direct approach. Moreover, it seems likely that the horizontal perturbations may require longer simulations for convergence than the vertical ones in Figure 2 because the latter involve only a single ligand, which normally binds in a single conformation, whereas the binding mode may change when the ligands are changed. In addition, the computational effort of the reference-potential approach can be further reduced by observing that the MM → QM/MM perturbation is performed for a single ligand, which shows that if the same ligand is involved in several perturbations, only the MM FEP needs to be redone, not the MM → QM/MM step for the common ligand. In the present case, this means that we could have omitted one MM → QM/MM perturbation for MeBz as well as three each for the Bz and Hx ligands, i.e., avoiding seven out of 16 MM → QM/MM perturbations. Thus, in total the reference-potential approach would be over 5 times more effective for the present test set.

The reference-potential method should also be compared to the ssEA approach used in our previous study.[29] We have used a total simulation time of 1.5 ns for each of the 12 MM → QM/MM perturbations in this study, each requiring 1.5 million QM/MM energy evaluations (with a time step of 1 fs). Thus, the present approach is ∼25 times more time-consuming than the ssEA approach used in our previous study.[29] On the other hand, the previous study involved at least two approximations with potentially serious influence on the results. First, it was essential to employ the cumulant approximation to calculate the ssEA MM → QM/MM free energies. This requires that the energy distribution is Gaussian, which is hard to prove for the low-energy tail that determines the free energy difference. In fact, even if the bootstrapped standard errors indicated that the free energies were converged to within 1 kJ/mol, the individual $\Delta\Delta G_L^{s,MM\rightarrow QM/MM}$ terms still showed differences of up to 9 kJ/mol when taken from independent simulations. In the present study, the corresponding terms were converged to better than 1 kJ/mol. Second, the previous study employed only interaction energies (E(complex) − E(host) − E(ligand)) in the MM → QM/MM calculations in order to improve the convergence. This excludes differences in the internal energies of the host and the ligand and may affect the results in an uncontrolled manner.[33] Finally, the present approach is fully automatic, employing only standard QM/MM software, whereas the previous approach required separate QM calculations of tens of thousands of structures, giving very many files and much overhead. For all of these reasons, we tend to prefer the present reference-potential approach.

Even if the present results are the best QM/MM results obtained for this system, it is of course somewhat disappointing that we still do not get any consistent improvement over the original MM calculations. However, this was not the aim of the present study. Instead, the QM method and size of the QM system (PM6-DH+ only for the ligand) were selected to allow for long and converged calculations and testing of several different approaches. If the prime aim is to improve the results, a more accurate method (e.g., dispersion-corrected DFT) should be used for the ligand as well as the host and the closest water molecules. The present study shows how such an investigation should be performed and allows for an estimate of the time consumption.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b01217.

The $\Delta G_L^{bound,MM\rightarrow QM/MM}$ and $\Delta G_L^{free,MM\rightarrow QM/MM}$ QM/MM free energy corrections for the eight transformations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Ulf.Ryde@teokem.lu.se. Tel: +46−46 2224502. Fax: +46−46 2228648.

ORCID

Ulf Ryde: 0000-0001-7653-8489

Funding

This investigation was supported by grants from the Swedish Research Council (Project 2014-5540) and the 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.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Gohlke, H.; Klebe, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644−2676.
(2) Gilson, M. K.; Zhou, H.-X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21−42.
(3) Coupez, B.; Lewis, R. A. *Curr. Med. Chem.* **2006**, *13*, 2995−3003.
(4) Kontoyianni, M.; Madhav, P.; Suchanek, E.; Seibel, W. *Curr. Med. Chem.* **2008**, *15*, 107−116.

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

(5) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(6) Genheden, S.; Ryde, U. *Expert Opin. Drug Discovery* **2015**, *10*, 449−461.
(7) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng., Des. Sel.* **1994**, *7*, 385−391.
(8) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 393−407.
(9) Hansen, N.; Van Gunsteren, W. F. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.
(10) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(11) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(12) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.
(13) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107.
(14) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *J. Med. Chem.* **2008**, *51*, 769−779.
(15) Christ, C.; Fox, T. *J. Chem. Inf. Model.* **2014**, *54*, 108−120.
(16) Mikulskis, P.; Genheden, S.; Ryde, U. *J. Chem. Inf. Model.* **2014**, *54*, 2794−2806.
(17) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. *J. Am. Chem. Soc.* **2015**, *137*, 2695−2703.
(18) Woo, H.-J.; Roux, B. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6825−6830.
(19) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 084108.
(20) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. *Chem. Sci.* **2016**, *7*, 207−218.
(21) Ryde, U.; Söderhjelm, P. *Chem. Rev.* **2016**, *116*, 5520−5566.
(22) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296−9297.
(23) Rathore, R. S.; Reddy, R. N.; Kondapi, A. K.; Reddanna, P.; Reddy, M. R. *Theor. Chem. Acc.* **2012**, *131*, 1096.
(24) Świderek, K.; Martí, S.; Moliner, V. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12614−12624.
(25) Luzhkov, V. B.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199−213.
(26) Duarte, F.; Amrein, B. A.; Blaha-Nelson, D.; Kamerlin, S. C. L. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 954−965.
(27) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94*, 138302.
(28) Mikulskis, P.; Cioloboc, D.; Andrejić, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Söderhjelm, P.; Ryde, U. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 375−400.
(29) Olsson, M. A.; Söderhjelm, P.; Ryde, U. *J. Comput. Chem.* **2016**, *37*, 1589−1600.
(30) Beierlein, F. R.; Michel, J.; Essex, J. W. *J. Phys. Chem. B* **2011**, *115*, 4911−4926.
(31) Genheden, S.; Ryde, U.; Söderhjelm, P. *J. Comput. Chem.* **2015**, *36*, 2114−2124.
(32) König, G.; Boresch, S. *J. Comput. Chem.* **2011**, *32*, 1082−1090.
(33) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. *J. Chem. Theory Comput.* **2016**, *12*, 499−511.
(34) Dybeck, E. C.; König, G.; Brooks, B. R.; Shirts, M. R. *J. Chem. Theory Comput.* **2016**, *12*, 1466−1480.
(35) Takahashi, H.; Omi, A.; Morita, A.; Matubayasi, N. *J. Chem. Phys.* **2012**, *136*, 214503.
(36) Sampson, C.; Fox, T.; Tautermann, C. S.; Woods, C. J.; Skylaris, C.-K. *J. Phys. Chem. B* **2015**, *119*, 7030−7040.
(37) Hummer, G.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 2004.
(38) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J. Chem. Theory Comput.* **2006**, *2*, 452−461.
(39) Woods, C. J.; Manby, F. R.; Mulholland, A. J. *J. Chem. Phys.* **2008**, *128*, 014109.
(40) Woods, C. J.; Shaw, K. E.; Mulholland, A. J. *J. Phys. Chem. B* **2015**, *119*, 997−1001.
(41) Hudson, P. S.; Woodcock, H. L.; Boresch, S. *J. Phys. Chem. Lett.* **2015**, *6*, 4850−4856.
(42) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B* **2011**, *115*, 7950−7962.
(43) Plotnikov, N. V.; Warshel, A. *J. Phys. Chem. B* **2012**, *116*, 10342−10356.
(44) Lameira, J.; Kupchencko, I.; Warshel, A. *J. Phys. Chem. B* **2016**, *120*, 2155−2164.
(45) Gibb, C. L. D.; Gibb, B. C. *J. Am. Chem. Soc.* **2004**, *126*, 11408−11409.
(46) Sun, H.; Gibb, C. L. D.; Gibb, B. C. *Supramol. Chem.* **2008**, *20*, 141−147.
(47) Gibb, C. L. D.; Gibb, B. C. *Tetrahedron* **2009**, *65*, 7240−7248.
(48) Muddana, H. S.; Fenley, A. T.; Mobley, D. L.; Gilson, M. K. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 305−317.
(49) Gibb, C. L. D.; Gibb, B. C. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 319−325.
(50) Andrejić, M.; Ryde, U.; Mata, R. A.; Söderhjelm, P. *ChemPhysChem* **2014**, *15*, 3270−3281.
(51) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173−1213.
(52) Korth, M. *J. Chem. Theory Comput.* **2010**, *6*, 3808−3816.
(53) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555−569.
(54) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(55) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(57) Tembre, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281−283.
(58) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A. E.; Salomon-Ferrer, R.; Simmerling, C.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, 2014.
(59) Walker, R. C.; Crowley, M. F.; Case, D. A. *J. Comput. Chem.* **2008**, *29*, 1019−1031.
(60) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(61) Wu, X.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381*, 512−518.
(62) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(63) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(64) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(65) Bhattacharyya, A. *Bull. Calcutta Math. Soc.* **1943**, *35*, 99−109.
(66) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.
(67) Genheden, S.; Ryde, U. *J. Comput. Chem.* **2010**, *31*, 837−846.

DOI: 10.1021/acs.jctc.6b01217
J. Chem. Theory Comput. 2017, 13, 2245−2253

Paper III

Casper Steinmann, Martin A. Olsson, and Ulf Ryde, *Journal of Chemical Theory and Computation, submitted,* **2018**

Relative Ligand-Binding Free Energies Calculated from Multiple Short QM/MM MD Simulations

Casper Steinmann,*,†,‡ Martin A. Olsson,‡ and Ulf Ryde*,‡

†*Department of Chemistry and Bioscience, Aalborg University, Frederik Bajers Vej 7H, DK-9220 Aalborg, Denmark*

‡*Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden*

E-mail: css@bio.aau.dk; ulf.ryde@teokem.lu.se

Abstract

We have devised a new efficient approach to compute combined quantum mechanical (QM) and molecular mechanical (MM, i.e. QM/MM) ligand-binding relative free energies. Our method employs the reference-potential approach with free-energy perturbation both at the MM level (between the two ligands) and from MM to QM/MM (for each ligand). To ensure that converged results are obtained for the MM→QM/MM perturbations, explicit QM/MM molecular dynamics (MD) simulations are performed with two intermediate mixed states. To speed up the calculations, we utilize the fact that the phase space can be extensively sampled at the MM level. Therefore, we run many short QM/MM MD simulations started from snapshots of the MM simulations, instead of a single long simulation. As a test case, we study the binding of nine cyclic carboxylate ligands to the octa-acid deep cavitand. Only the ligand is in the QM system, treated with the semi-empirical PM6-DH+ method. We show that for eight of the ligands, we obtain well converged results with short MD simulations (1–15 ps). However, in one case, the convergence is slower (∼50 ps) owing to a mismatch between the

conformational preferences of the MM and QM/MM potentials. We test the effect of initial minimization, the need of equilibration and how many independent simulations are needed to reach a certain precision. The results show that the present approach is about 5 times faster than using standard MM→QM/MM free-energy perturbations with the same accuracy and precision.

1 Introduction

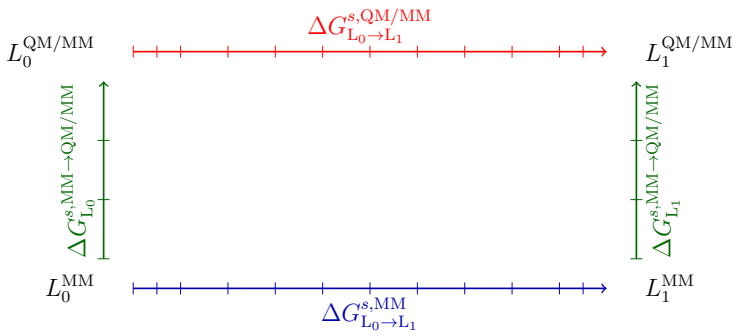
Predicting the binding of a small molecule (L for ligand) to a biological macromolecule (R for receptor), i.e. the free energy of the reaction R + L → RL, remains one of the greatest challenges of computational chemistry. Accurate predictions of free energies could have a large impact on areas such as rational drug design, for which the need to perform expensive experiments would be reduced.

According to statistical mechanics, the free-energy difference can be expressed as averages over molecular configurations of the atomic coordinates, which can be obtained from molecular dynamics (MD) or Monte Carlo simulations. In principle, proper free energies can be obtained from free-energy perturbation (FEP) calculations with the energies estimated by exponential averaging (EA, Zwanzig equation),[1] thermodynamic integration,[2] Bennett acceptance ratio (BAR)[3,4] or similar methods. The difference in binding free energies of two similar ligands (L_0 and L_1) can be obtained by calculating the energy difference between the two ligands when bound to the receptor macromolecule and when free in solution.[5] However, such calculations converge only when the energy difference is small, i.e. when there is enough overlap in the energy functions of the two ligands so that they sample the same part of the phase space. In practice, this is seldom the case and this problem is solved by dividing the transformation between the two ligands into several small steps, employing an interpolated potential, using a coupling parameter λ ($0 \leq \lambda \leq 1$):

$$E(\lambda) = (1 - \lambda)E_0 + \lambda E_1, \tag{1}$$

where E_0 and E_1 are the potentials for L_0 and L_1, respectively. Moreover, FEP is guaranteed to yield correct results only if the energy function is perfect and the sampling is exhaustive. The latter point has historically implied the use of molecular mechanics (MM) force fields for the sampling, because such calculations are cheap enough to allow for long simulation times that provide a representative ensemble of configurations.[6,7]

However, the MM force field is far from a perfect energy function and therefore, there has been quite some interest to use quantum mechanical (QM) methods in the FEP simulations.[8] For the binding of a ligand to a bio-macromolecule receptor, this would require the use of combined QM/MM methods.[9,10] Such calculations can be performed at many levels of approximations, as is illustrated in Scheme 1. The arrow at the top of the figure, represents



Scheme 1: Thermodynamic cycle employed to estimate the relative binding free energy between two ligands (L_0 and L_1) at the QM/MM level of theory, $\Delta G^{s,\mathrm{QM/MM}}_{\mathrm{L_0\rightarrow L_1}}$, for state s (i.e. either for the ligand free in solution or when bound to the host). In the reference-potential methods, $\Delta G^{s,\mathrm{QM/MM}}_{\mathrm{L_0\rightarrow L_1}}$ is estimated from the change in free energy at the MM level of theory, $\Delta G^{s,\mathrm{MM}}_{\mathrm{L_0\rightarrow L_1}}$ (lower horizontal arrow). The free energies are corrected at the end points with a FEP in methods space, from MM to QM/MM, $\Delta G^{s,\mathrm{MM\rightarrow QM/MM}}_{\mathrm{L_i}}$. The bars on the arrows illustrate that each perturbation is divided into several intermediate states, characterized by the λ coupling parameter for the horizontal perturbations and by the Λ parameter for the vertical perturbations.

a full FEP calculation at the QM/MM level, $\Delta G^{s,\mathrm{QM/MM}}_{\mathrm{L_0\rightarrow L_1}}$, where s, represents either the ligand bound to the receptor or free in solution. The net binding free energy, $\Delta\Delta G^{\mathrm{bind}}_{\mathrm{L_0\rightarrow L_1}}$, is

the difference of those two terms:

$$\Delta\Delta G^{\mathrm{bind,QM/MM}}_{\mathrm{L_0\to L_1}} = \Delta G^{\mathrm{bound,QM/MM}}_{\mathrm{L_0\to L_1}} - \Delta G^{\mathrm{free,QM/MM}}_{\mathrm{L_0\to L_1}}. \quad (2)$$

Such FEP calculations have been employed a few times for protein binding energies,[11–14] using a semi-empirical QM method and treating only the ligand by QM. The reason for this is that long MD simulations and many intermediate states are required to get converged and reliable results (for example we run $2 \times 18 \times 1.5$ ns $= 54$ ns simulations for each ligand pair in our recent study[14]). This has spurred the development of computationally cheaper alternatives, in particular the reference-potential methods.[15–18] These take advantage of the corresponding FEP calculations at the MM level ($\Delta G^{s,\mathrm{MM}}_{\mathrm{L_0\to L_1}}$ at the bottom of Scheme 1), which are much cheaper. By using the thermodynamic cycle in Scheme 1, the $\Delta G^{s,\mathrm{QM/MM}}_{\mathrm{L_0\to L_1}}$ free energy can be obtained from $\Delta G^{s,\mathrm{MM}}_{\mathrm{L_0\to L_1}}$ by performing FEP simulations in method space to convert the MM potential to a QM/MM (or QM) potential ($\Delta G^{s,\mathrm{MM\to QM/MM}}_{\mathrm{L_i}}$ for the two ligands L_0 and L_1 shown as vertical arrows in Scheme 1), i.e.

$$\Delta\Delta G^{\mathrm{bind,QM/MM}}_{\mathrm{L_0\to L_1}} = \Delta\Delta G^{\mathrm{bind,MM}}_{\mathrm{L_0\to L_1}} - \Delta\Delta G^{\mathrm{bind,MM\to QM/MM}}_{\mathrm{L_0}} + \Delta\Delta G^{\mathrm{bind,MM\to QM/MM}}_{\mathrm{L_1}}. \quad (3)$$

Again, several approaches can be used to calculate $\Delta G^{s,\mathrm{MM\to QM/MM}}_{\mathrm{L_i}}$. The most common approach is to try to estimate it by exponential averaging. If this converges in a single step (ssEA), no MD simulations at the QM/MM level of theory are needed because the energies can be calculated directly by single-point QM calculations on snapshots already available from the FEP calculation of $\Delta G^{s,\mathrm{MM}}_{L_0\to L_1}$.[15,16,19] Alternatively, $\Delta G^{s,\mathrm{MM\to QM/MM}}_{\mathrm{L_i}}$ can be calculated by the non-Boltzmann BAR approach (NBB), which employs the fact that BAR normally gives better results than EA, especially when the overlap of the energy functions is poor.[20] However, NBB requires QM calculations on an additional intermediate state and therefore is twice as expensive as ssEA.[19] Moreover, theoretical considerations have indicated that ssEA performs slightly better than NBB both in theory and in practice.[21]

NBB and ssEA have been used in recent studies of binding free energies.[21–25] However, several of these (especially when including more than the ligand in the QM system[23,24]) showed serious problems to converge the free energies, simply because the MM and QM/MM potentials are too dissimilar to allow $\Delta G^{s,\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{L_i}}$ to be calculated based solely on snapshots from MD simulations at the MM level. For enzyme reaction energies, the problem is often solved by keeping the QM system fixed in the FEP calculations,[15–17] but that does not seem to be a useful approach for ligand binding, for which entropy effects of the ligand are important. With a semi-empirical QM method, we have managed to obtain $\Delta G^{s,\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{L_i}}$ energies that were seemingly converged to within 1 kJ mol^{-1} employing 720 000 QM calculations per ligand, but only when using ssEA with the cumulant approximation to the second order[26,27] (ssEAc; i.e. assuming that the energy distribution is Gaussian).[19] Plain ssEA and NBB gave results with an estimated uncertainty of 2–7 kJ mol^{-1}.

This number of QM/MM energy calculations corresponds to up to 1.4 ns of MD simulations. Therefore, we tested in our most recent study to actually run QM/MM MD simulations.[14] We performed both full QM/MM FEP calculations (upper arrow in Scheme 1) and reference-potential calculations, but dividing the calculations of $\Delta G^{s,\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{L_i}}$ (vertical arrows in Scheme 1) into several steps denoted by the coupling parameter Λ for the hybrid energy function

$$E(\Lambda) = (1-\Lambda)E_{\mathrm{MM}} + \Lambda E_{\mathrm{QM/MM}}, \tag{4}$$

connecting the MM energy (E_{MM}) and the QM/MM energy ($E_{\mathrm{QM/MM}}$; again $0 \leq \Lambda \leq 1$); we use Λ to avoid confusion with the λ parameter for the horizontal perturbations in Scheme 1). Thus, the calculation of $\Delta G^{s,\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{L_i}}$ employed expensive QM/MM MD simulations. We showed that the two approaches gave identical results, within the statistical uncertainty. For the direct calculation of $\Delta G^{s,\mathrm{QM/MM}}_{\mathrm{L_0}\rightarrow\mathrm{L_1}}$, our overlap measures indicated that a total of 18 λ values were needed. For $\Delta G^{s,\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{L_i}}$, it was enough to consider four Λ values. Consequently, the latter approach was more effective. However, since each state required 1.5 ns MD simulations, this approach was much more demanding than the ssEAc calculations.

On the other hand, ssEAc depends on the accuracy of the cumulant approximation and it is normally performed only for interaction energies and not for the more correct total energies (again to improve convergence). Moreover, there were indications that the results were still not fully converged with independent calculations of the same energy component yielding results that varied by up to 9 kJ/mol.[19] Therefore, the reference-potential calculations with explicit QM/MM MD simulations seemed to be more reliable. Other studies have also showed that converged binding affinities can be obtained with sampling at the QM/MM level.[25,28] A similar approach (paradynamics) has been used for enzyme reactions, but normally with only two Λ states[15,29,30]), but for our binding affinities, this gave poor overlap and not converged results.[14] Another study has recently reached similar conclusions.[31]

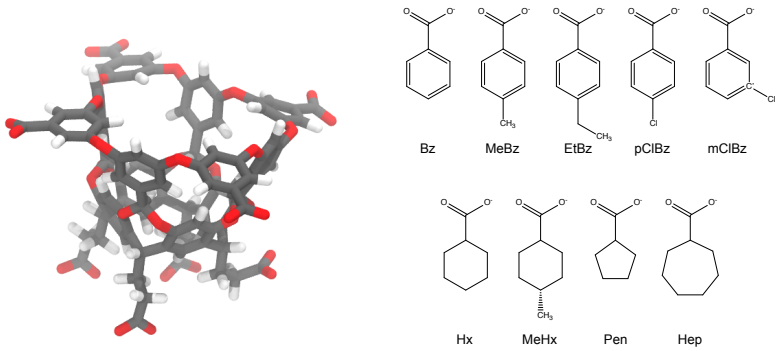
In this paper, we use a similar method, i.e. the reference-potential approach, calculating $\Delta G_{\mathrm{L_i}}^{\mathrm{s,MM\rightarrow QM/MM}}$ with four Λ states, but investigate whether the simulations can be sped up, exploiting the fact that we can thoroughly explore the phase space for each ligand with MD simulations at the MM level. Therefore, we make the QM/MM MD simulations very short and instead run many of them. We investigate if we still can reproduce the previous results and examine how long the simulations need to be, how much equilibrium time is needed and how many independent calculations are needed to reach a certain precision. The results show that significant time can be saved by such an approach without compromising the accuracy for most systems.

2 Computational Details

2.1 Ligands and MM calculations

We have studied the MM→QM/MM contribution to the binding free energies for nine cyclic carboxylate ligands bound to the octa-acid (OA) deep-cavity cavitand from the SAMPL4[32,33] competition, shown in Scheme 2. The aim is to improve the estimates of the relative binding free energies in the following eight ligand transformations: MeBz→Bz, EtBz→MeBz,

pClBz→Bz, mClBz→Bz, Hx→Bz, MeHx→Hx, Hx→Pen and Hx→Hep (the names of the ligands are defined in Scheme 2). The underlying MM simulations and estimates of the



Scheme 2: Octa-acid host (left) and ligands considered in this study (right).

relative binding free energy at the MM level of theory were taken from our previous work on the same system.[19,23] In those, the general Amber force field[34] (GAFF) was used for both the host and the ligands, with partial atomic charges derived from restrained electrostatic potential fits[35] (RESP), obtained at the HF/6-31G* level of theory.

2.2 QM/MM Free-Energy Simulations

For each ligand, either bound to the host or free in water solution, we extracted 100 snapshots from the previous MD simulations at the MM level. For each of these snapshots, we ran a 50 ps QM/MM FEP calculation using four Λ values ($\Lambda = 0.0, 0.333, 0.666, 1.0$).[14] The QM/MM MD simulations were done with the *sander* module of the Amber 14[36] software using the dual-topology scheme. Only the ligand was included in the QM region, whereas the host and the solvent were in the MM system. The QM calculations were performed at the semi-empirical PM6-DH+[37] level of theory. In all calculations, the temperature was kept at 300 K using Langevin dynamics with a collision frequency of 2 ps^{-1}.[38] The pressure was kept at 1 atm using Berendsen's weak-coupling isotropic algorithm with a relaxation time of 1 ps.[39] Long-range electrostatics were handled by particle-mesh Ewald summation,[40] whereas

Lennard-Jones interactions were truncated beyond 8 Å. All QM/MM MD simulations were run without any restraints and with a time step of 1 fs.

For each snapshot, free-energy differences were calculated over the four Λ-windows with the multistate Bennett acceptance ratio (MBAR) approach, as implemented in the pyMBAR software.[41]

3 Results and Discussion

In this paper, we studied whether calculations of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ with the reference-potential approach can be sped up by using many short QM/MM simulations, exploiting the fact that the cheap MM simulations already have extensively explored the phase space. To this aim, we have studied the binding of nine small cyclic carboxylate molecules (shown in Scheme 2) to the octa-acid deep-cavity host. We have already studied this system in several previous studies, giving us proper reference results.[14,19,23] In separate subsections, we will describe first the convergence of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ with respect to the simulation time, then the number of independent simulations needed and finally, whether the results can be improved by omitting an initial part of the simulation as an equilibration period.

3.1 Convergence of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$

In this section, we discuss how the MM→QM/MM free-energy change for a ligand $\mathrm{L_i}$, $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$, converges as a function of simulation length for the nine ligands in this study. The free-energy changes were calculated using MBAR at simulation lengths (τ) ranging from 1 to 50 ps. We computed the mean of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ and standard error of the mean ($\mathrm{SE_{L_i}} = s/\sqrt{N}$ where s is standard deviation) over the N trajectories included in the averaging procedure ($N = 100$ if not otherwise stated). These results are presented in Figure 1, in which we also show the $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ values obtained from our previous study with full QM/MM FEP simulations (Table 1 in Ref 14; 1.5 ns MD simulations for

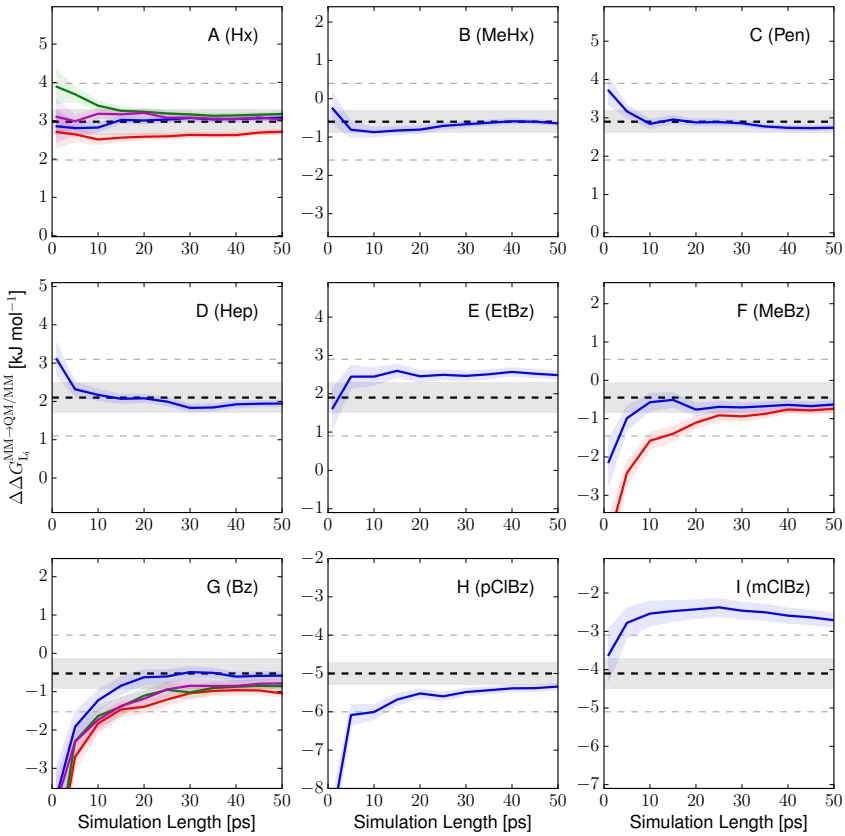


Figure 1: Convergence profiles of $\Delta\Delta G_{L_i}^{MM\rightarrow QM/MM}$ for the nine ligands in this study as a function of the simulation time per window.

four Λ values) as black dashed lines with uncertainties plotted as gray shaded areas. These results are used as reference values for the new (shorter) simulations. We have included markers (dashed gray horizontal lines) at ± 1 kJ mol^{-1} from the reference value to indicate a satisfactory accuracy, in line with previous efforts to quantify convergence.[14,19] It also represents an approximate confidence interval for the difference between the old results

(standard errors of 0.3–0.4 kJ mol^{-1}) and new results (standard errors decreasing from 0.4 to 0.1 kJ mol^{-1} when the simulation time is increased, employing a t-factor of 1.96 (95% confidence). We emphasize that the time shown on the x-axis in Figure 1 is given in picoseconds and represents the simulation time for each Λ window. This serves to elucidate how fast $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ converges with the length of the simulations. QM/MM minimizations prior to QM/MM MD simulations had negligible effects on the convergence (see Figure S1 in the Supporting Information) and was therefore not further investigated. **CS:**Figure S1 min10000

The results show that for five of the ligands (Hx, MeHx, Pen, Hep and EtBz shown in panels A to E in Figure 1), $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ is converged to within 1 kJ mol^{-1} of the reference value already from the first picosecond of the simulation. For $\Delta\Delta G_{\mathrm{Hx}}^{\mathrm{MM\rightarrow QM/MM}}$, we have results from four independent sets of simulations (taken from the transformations Hx→Bz, MeHx→Hx, Hx→Pen and Hep→Hx) and these are shown in different colors in Figure 1A. It is satisfying that the four simulations give results that agree with each other within 1 kJ mol^{-1} already from 1 ps simulation time. This is actually better than in the previous simulations, in which there was a variation of 1.5 kJ mol^{-1} (2.1–3.6 kJ mol^{-1}) and nine Λ values were needed to bring the variation down to 0.3 kJ mol^{-1}. For four of the ligands, Hx, MeHx, Pen, Hep, the new results agree with the reference values within 0.3 kJ mol^{-1} at the end of the simulation (50 ps), and actually already after 1–20 ps simulation time. However, for EtBz, there is a 0.6 kJ mol^{-1} difference between the new results (apparently converged already after 5 ps) and the reference value (Figure 1E). In principle, this is no problem, as the new result still falls within the 95% (and also 90%) confidence interval of the reference results (the gray area in the figure marks a single standard deviation). Yet, considering the good accuracy and precision of the other new calculations and the clear convergence of the results, it seems likely that the new result is more accurate. This is confirmed by the fact that in the previous investigation, nine Λ values gave a slightly larger $\Delta\Delta G_{\mathrm{EtBz}}^{\mathrm{MM\rightarrow QM/MM}}$, 2.1±0.5 kJ mol^{-1},[14] closer to the new result.

For three of the ligands, MeBz, Bz and pClBz, it is slightly more difficult to converge $\Delta\Delta G_{\mathrm{Li}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ to within 1 kJ mol^{-1} of the reference values. For $\Delta\Delta G_{\mathrm{MeBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$, shown in Figure 1F, convergence is achieved after 3 and 10 ps for the two independent calculations (based on results for the MeBz→Bz and EtBz→MeBz transformations) and the final results (at 50 ps) agree closely with each other and with the reference value. The free energy change for the Bz ligand, $\Delta\Delta G_{\mathrm{Bz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$, shown in Figure 1G, has slightly longer convergence times ranging from 9 to 15 ps for the four independent sets of simulations (taken from the transformations MeBz→Bz, pClBz→Bz, mClBz→Bz and Hx→Bz). The pClBz ligand shows similar convergence, in that it requires roughly 10 ps of simulation time to converge $\Delta\Delta G_{\mathrm{pClBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$, despite deviating by more than 3 kJ mol^{-1} initially, as shown in Figure 1H.

However, the results for the mClBz ligand is much different from those of the other eight ligands: Even after 50 ps, $\Delta\Delta G_{\mathrm{mClBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ deviates by 1.4 kJ mol^{-1} from the reference value and the results do not seem to be converged. $\Delta\Delta G_{\mathrm{mClBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ attains a maximum (–2.2 kJ mol^{-1}) at 25 ps and then slowly decreases during the remainder of the simulation time. Therefore, we extended the simulation time to 150 ps and the results are shown in

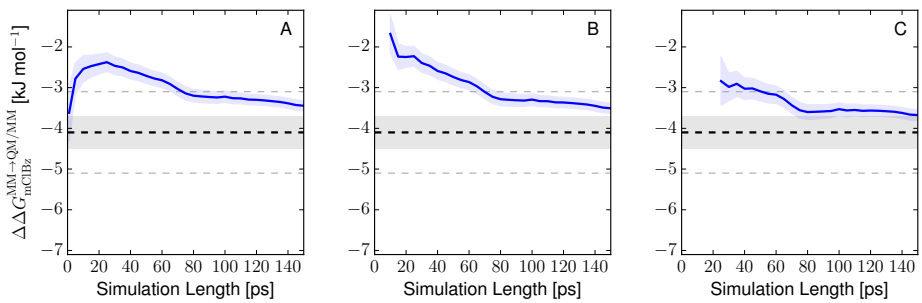


Figure 2: Results of extended (150 ps) simulations for the mClBz ligand. In A), all data were employed, whereas in B) and C) the first 10 ps and 25 ps were omitted in the estimate of $\Delta\Delta G_{\mathrm{mClBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$.

Figure 2A, in which it can be seen that the results agrees within 1 kJ mol^{-1} of the reference

value after 75 ps, but the curve is still decreasing at the end of the simulation. The results do not change much if the first 10 ps of the simulation is discarded as an equilibration period, besides that the maximum is increased and moves to 10 ps (Figure 2B). However, if the equilibration period is extended to 25 ps (Figure 2C), convergence to within 1 kJ mol^{-1} is reached already after 50 ps simulation time and $\Delta\Delta G^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{mClBz}}$ converges to a nearly constant value after 80 ps time. This indicates that mClBz undergoes a slow conformational change, giving less negative $\Delta\Delta G^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}_{\mathrm{mClBz}}$ at the beginning of the simulation, which takes a very long time to average away.

A thorough investigation of the energies and the simulations (described in the Supporting Information) showed that the slow convergence can be traced to the simulation of the complex and in particular the position of the Cl atom in the host. To illustrate this difference, we defined the vector between the centroid of the upper and lower ring in the octa-acid host, $\mathbf{R}_{\mathrm{OA}}$, shown in Figure S12. Then, the angle between $\mathbf{R}_{\mathrm{OA}}$ and the C'–Cl vector (see Scheme 2) in mClBz was followed during the MD simulations. From the histograms in Figure 3, it can be seen that at 10 ps simulation time, there is only a minor difference in the angles attained with the MM and QM/MM potentials, both showing a small peak around ∼30° (Cl atom points almost downwards in the host) and a larger and broader peak at ∼60° (Cl atom points towards the side of the host). However, as the simulation time is increased, a shift in

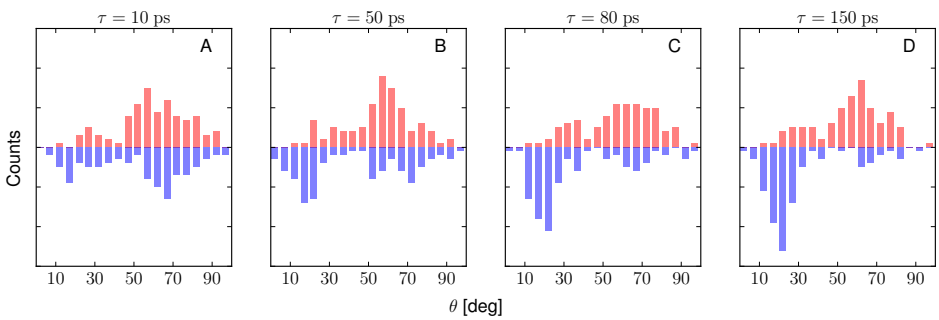


Figure 3: Histograms of the angle between the host $\mathbf{R}_{\mathrm{OA}}$ and mClBz ligand C'–Cl vectors at different simulation times. The upper (red) histograms are for the MM potential at $\Lambda = 0.0$ and the blue histograms are for the QM/MM potential at $\Lambda = 1.0$.

the angles obtained with the QM/MM potential is observed towards $< 30°$ (Figure 3B–D), whereas the angles for the MM potential do not appear to change significantly. From the results presented in Figure 3, it is clear that this conformational change takes a long time ($\sim$50 ps) and it is the reason for the slow convergence observed in Figures 1 and 2. The problem probably stems from sub-optimal parameters of the chlorine atom in the GAFF force field of mClBz and may be ameliorated by a re-parametrization of the van der Waals parameters.

In summary we find that the MM→QM/MM calculations using our approach works well for eight of nine ligands, with convergence of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ within 1–15 ps, considerably reducing the cost of the simulations. However, for a single system, a significant difference between the conformations attained by the ligand in the MM and QM/MM potential is observed and a proper equilibration of this difference takes $\sim$50 ps.

3.2 Effect of Equilibration Period

For the mClBz ligand, it was seen in Figure 2 that the results were improved if the initial 25 ps simulations were discarded as an equilibrium period. It is possible that also the other results could be improved by such a procedure. Therefore, we have tested to discard the first 1 or 5 ps of the simulations. The results are shown in Figures S13 and S14 in the Supporting Information. It is seen that for MeBz, Bz and pClBz, the results strongly improve if some data from the first part of the simulation is discarded. If 1 ps is discarded, the pClBz simulation is converged already from start, whereas $\Delta\Delta G_{\mathrm{MeBz}}^{\mathrm{MM\rightarrow QM/MM}}$ converges within 3–7 ps (4–13 ps without any equilibration) and $\Delta\Delta G_{\mathrm{Bz}}^{\mathrm{MM\rightarrow QM/MM}}$ converges within 4–7 ps, which is an even larger improvement (8–14 ps without any equilibration). With 5 ps equilibration, all seven free energies are converged already at the start of the sampling period (5 ps). For none of the ligands, the final result (at 50 ps) changes if data is discarded as equilibration.

The following systematic approach can be used to determine how much time (t_{eq}) at the start of the simulation should be discarded as equilibration. We define as our refer-

ence $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ (with standard error $\mathrm{SE}_{\mathrm{L_i}}$) computed as the result obtained by integrating from t_{eq} to the end of the simulation ($t_{\mathrm{sim}} = 50$ ps) using the following notation $\Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{sim}})$ for the free energy and $\mathrm{SE}_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{sim}})$ for the standard error (omitting the superscript MM→QM/MM for simplicity). We then test if the first picosecond of the simulation can be discarded based on whether it deviates significantly from the reference results by a simple t-test: We compute $\Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{eq}} + 1\text{ ps})$ and $\mathrm{SE}_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{eq}} + 1\text{ ps})$ and measure whether these two free energy distributions overlap by comparing the absolute difference in free energies and the composite error of the two standard errors with 95 % confidence following Nicholls:[42]

$$|\Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{sim}}) - \Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{eq}} + 1\text{ ps})| < t_{95\%} \cdot \mathrm{CE}_{\mathrm{L_i}}. \tag{5}$$

where the composite error for ligand $\mathrm{L_i}$, $\mathrm{CE}_{\mathrm{L_i}}$, is

$$\mathrm{CE}_{\mathrm{L_i}} = \sqrt{\left(\mathrm{SE}_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{sim}})\right)^2 + \left(\mathrm{SE}_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{eq}} + 1\text{ ps})\right)^2}$$

If the inequality in eq 5 does not hold, the first picosecond is discarded as equilibration time so $t_{\mathrm{eq}} = t_{\mathrm{eq}} + 1$ ps and both $\Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{sim}})$ and $\Delta\Delta G_{\mathrm{L_i}}(t_{\mathrm{eq}}, t_{\mathrm{eq}} + 1\text{ ps})$ are recomputed and eq 5 is re-evaluated until it holds. This procedure was applied to all ligands (except the diverging mClBz ligand) for $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$, but also on the individual $\Delta G_{\mathrm{L_i}}^{s,\mathrm{MM\rightarrow QM/MM}}$ values for the ligand free in solution or bound to the host. These results are presented in Table 1.

It can be seen that the equilibration time in general is shorter for $\Delta\Delta G$ than for ΔG^s. This is in agreement with the data presented in Table S1 and Figures S2–S10, discussed in section 2 of the SI. For $\Delta\Delta G$, t_{eq} is 1–3 ps for the aromatic ligands (5 ps for MeBz) and 0 ps for the non-aromatic ligands (1 ps for Hep). For ligands free in solution, the equilibration times are 2–6 ps and when bound to the octa-acid host they are 2–4 ps, except for one case each for MeBz (8 ps) and Hx (10 ps). Apart from the latter two outliers, the ligands bound

Table 1: Equilibration period, t_{eq} (in ps), for the various ligands, employing the results after 50 ps as the reference value. t_{eq} was calculated for either $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$, or individually for the simulations of the ligand bound to the host or free in solution. For the Bz, MeBz and Hx ligands, four or two independent results are presented, as has been explained before.

$\mathrm{L_i}$	$t_{\mathrm{eq}}(\Delta\Delta G)$	$t_{\mathrm{eq}}(\Delta G^{bound})$	$t_{\mathrm{eq}}(\Delta G^{free})$
Bz	2	3	4
	2	3	4
	3	3	4
	3	2	5
MeBz	2	8	4
	5	3	6
EtBz	1	4	4
pClBz	1	4	4
Hx	0	2	4
	0	2	6
	0	10	4
	0	2	5
MeHx	0	2	4
Pen	0	2	2
Hep	1	3	3

to the octa-acid host equilibrate faster than when they are free in solution. The longer equilibration times for ΔG^s is a result of the larger magnitude of these energies (500–800 kJ mol^{-1}).

3.3 Number of independent simulations

To thoroughly converge $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$, we have so far assumed that 100 snapshots are required. However, it is clear from the results in Figure 1 that the standard errors are very small ($\sim$0.1 kJ mol^{-1} for 50 ps simulation time), indicating that the number of snapshots may be reduced without compromising the accuracy. To test this, we performed bootstrap simulations with 5000 samples to estimate the standard error of the mean as a function of the number of included snapshots for three simulation times (10, 20 and 50 ps). The results are presented in Figure 4 for all nine ligands.

It can be seen that for all ligands except mClBz (which we already have discussed in

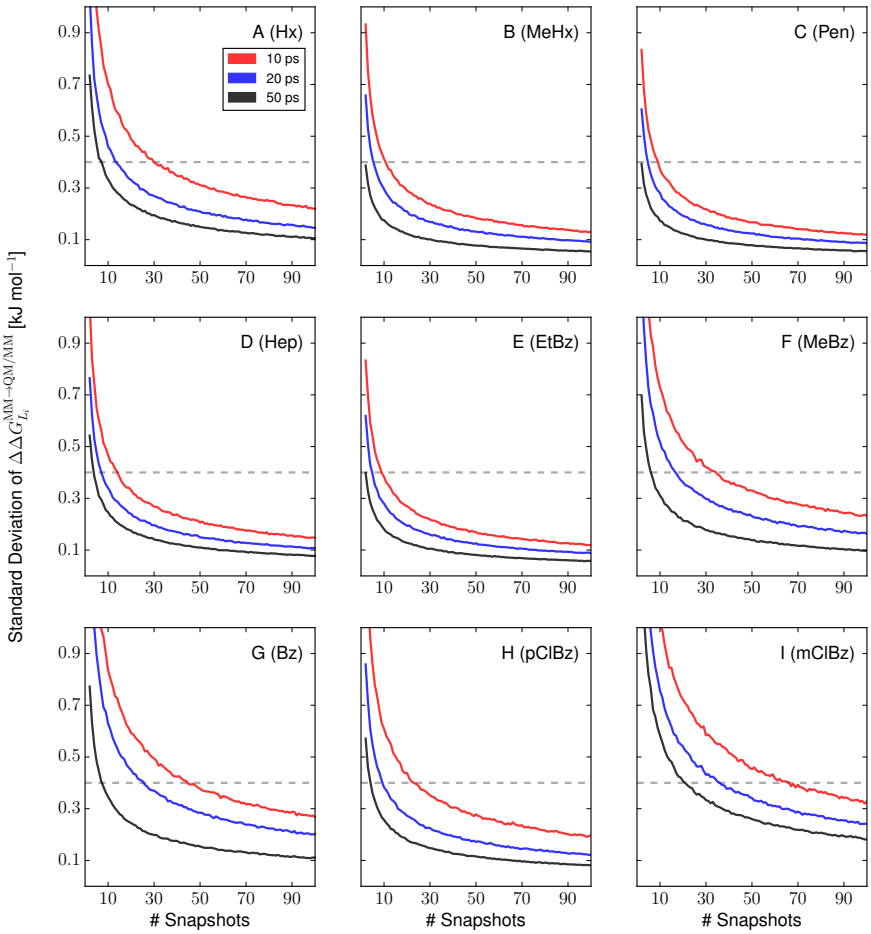


Figure 4: Simulated standard error of the free energy change $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ using bootstrapping.

detail), the standard error falls rapidly off for the full 50 ps simulations. Besides, mClBz, the Bz ligand shows the slowest convergence (Figure 4G), so we concentrate on this ligand. The precision in our previous study of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM\rightarrow QM/MM}}$ in this system was 0.4 kJ mol^{-1} for most ligands using four Λ values. This value is marked by a dashed horizontal line in Figure 4. For such a precision the Bz ligand requires 45, 25 and 8 snapshots using

simulations of lengths 10 ps, 20 ps and 50 ps, respectively. Based on these results, we suggest as a compromise between accuracy and computational effort to use 20 simulations of 20 ps of simulation time in each Λ-window (**OPT0**). In Table 2, $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ results for

Table 2: Computed $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ values for all nine ligands using various strategies (described in the text). Energies are in kJ mol^{-1}, equilibration times in ps. As usual, two or four independent results are given for Bz, MeBz and Hx.

L_i	**OPT0**	**OPT1**	n_s	**OPT2**	n_s	t_{eq}	$\Lambda = 4^a$	$\Lambda = 9^a$
Bz	-0.1 ± 0.5	-0.6 ± 0.3	43	-0.4 ± 0.3	43	1	-0.3 ± 0.4	-1.0 ± 0.5
	-1.1 ± 0.4	-1.7 ± 0.3	32	-1.5 ± 0.3	32	1	-0.3 ± 0.4	-0.9 ± 0.5
	-1.2 ± 0.4	-1.4 ± 0.3	33	-1.4 ± 0.3	33	1	-0.8 ± 0.4	-1.0 ± 0.5
	-1.7 ± 0.5	-1.4 ± 0.3	43	-1.3 ± 0.3	43	1	-0.7 ± 0.4	-0.8 ± 0.5
MeBz	-0.4 ± 0.4	-0.3 ± 0.3	18	-0.2 ± 0.3	18	3	-0.6 ± 0.4	-1.3 ± 0.5
	-1.0 ± 0.4	-0.9 ± 0.3	31	-0.7 ± 0.3	31	1	-0.3 ± 0.4	-0.8 ± 0.5
EtBz	2.6 ± 0.3	2.2 ± 0.3	24	2.3 ± 0.3	30	2	1.9 ± 0.4	2.1 ± 0.5
pClBz	-5.4 ± 0.3	-5.8 ± 0.3	18	-5.5 ± 0.3	19	2	-5.0 ± 0.3	-5.5 ± 0.5
Hx	3.1 ± 0.2	2.7 ± 0.3	7	2.7 ± 0.2	8	3	2.9 ± 0.3	2.9 ± 0.4
	2.5 ± 0.2	1.8 ± 0.3	5	1.8 ± 0.2	5	2	2.1 ± 0.4	2.7 ± 0.4
	3.0 ± 0.2	3.3 ± 0.3	17	3.2 ± 0.3	17	1	3.3 ± 0.3	2.7 ± 0.4
	3.0 ± 0.2	3.3 ± 0.3	5	3.2 ± 0.3	5	1	3.6 ± 0.3	3.0 ± 0.4
MeHx	-0.6 ± 0.3	-1.0 ± 0.3	13	-1.1 ± 0.3	13	1	-0.6 ± 0.3	-0.6 ± 0.4
Pen	3.0 ± 0.2	2.8 ± 0.3	6	2.8 ± 0.3	6	0	2.9 ± 0.3	2.4 ± 0.5
Hep	2.2 ± 0.3	2.1 ± 0.3	13	2.1 ± 0.3	13	0	2.1 ± 0.4	0.9 ± 0.4

[a] previous results taken from Ref 14.

OPT0 is presented (employing every fifth snapshot out of the original 100) for each ligand. In general, the results agree with reference results from our previous work[14] but obtained at almost a fifth of the computational cost (20 simulations of 20 ps correspond to a total simulation time of 400 ps, compared to 1.5 ns simulations in each Λ-window used previously). For example, the pClBz ligand is predicted to have a $\Delta\Delta G_{\mathrm{pClBz}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ value of -5.4 ± 0.3 kJ mol^{-1} which is statistically equivalent to -5.0 ± 0.3 kJ mol^{-1} and -5.5 ± 0.5 kJ mol^{-1} obtained previously using four and nine Λ-values, respectively. The mean absolute deviation is 0.4 and 0.5 kJ mol^{-1} from the results with four and nine Λ values, respectively, and the maximum deviation is 1.0 and 1.3 kJ mol^{-1}, always within the 95% confidence intervals. Not unexpectedly, using a fixed number of snapshot gives rise to a varying precision with standard errors varying between 0.5 kJ mol^{-1} for Bz and 0.2 kJ mol^{-1} for most of the

non-aromatic ligands, which is in agreement with the presented in Figure 4 above. We have again excluded the diverging mClBz ligand.

Alternatively, we may want to specify the desired precision of the results then perform the number of simulations needed to reach such a precision (this is easily obtained by first performing a few simulations, then calculating the standard error and use the expected square-root dependence to estimate the number of simulations needed). On the other hand, such an approach (**OPT1**) gives rise to a varying computational cost for the various ligands. In Table 2 results for **OPT1** are presented with a precision of 0.3 kJ mol^{-1} (still with 20 ps simulation time) together with the required number of snapshots, n_{s}. The results are very similar to the **OPT0** results, showing differences of up to 0.7 kJ mol^{-1} (for one of the calculations with the Hx ligand), but in general less than 0.3 kJ mol^{-1} for the other ligands. As expected, n_{s} was less than 20 for the non-aromatic ligands ($n_{\mathrm{s}} = 5$–17). For the aromatic ligands, n_{s} is between 18 (MeBz and pClBz) and 43 (Bz). The mean absolute deviations from the $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ results obtained with four and nine Λ values are again 0.4 and 0.5 kJ mol^{-1}, respectively, with maximum deviations of 1.4 and 1.2 kJ mol^{-1}.

We also incorporated our equilibration approach discussed in Section 3.2. We use the procedure outlined for **OPT1** to initially obtain $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ with a precision of 0.3 kJ mol^{-1}. This was followed by the equilibration procedure, in which initial parts of the simulation were discarded based on the inequality in eq 5. If the standard error increased by discarding the equilibration data, n_{s} was increased by one and the equilibration procedure was repeated. Computed values of $\Delta\Delta G_{\mathrm{L_i}}^{\mathrm{MM}\rightarrow\mathrm{QM/MM}}$ together with the number of snapshots n_{s} and equilibration times t_{eq} are also presented in Table 2 with this approach (**OPT2**). We find that the aromatic ligands require at least some equilibration (Bz, MeBz and pClBz has $t_{\mathrm{eq}} = 1$–3 ps) which is consistent with the results in obtained in Table 1. The computed mean absolute deviation was found to be 0.4 and 0.5 kJ mol^{-1}, respectively. The corresponding maximum deviations were 1.2 and 1.2 kJ mol^{-1}, respectively, quite similar to the other two procedures.

Finally, we note that it is also possible to estimate the convergence of the results through overlap measures. This was done thoroughly in our previous studies of the same and other systems.[14,19,23] The conclusion was that the Π bias metric by Wu and Kofke[43] gave the most reliable results. Therefore, we have investigated the Π value for all individual simulations. The results, presented in Figure S15, shows that all simulations give $\Pi > 0$, indicating a proper overlap between the two energy distributions (MM and QM/MM).[43] In fact, only five simulations give $\Pi < 0.5$, which has been suggested as a heuristic safety margin.[43] In principle, simulations with $\Pi < 0.5$ could be discarded, elongated or re-run using five or more Λ-windows but this was not tested in this work.

4 Conclusions

We have presented a new efficient approach to compute free energy at the QM/MM level with the reference-potential method, which shows promising convergence properties. This new approach uses extensive conformational sampling at the MM level combined with multiple short QM/MM MD simulations to compute free energies at the QM/MM level. Such a combination provides a fast way to converge MM→QM/MM free energies and requires ∼5 times shorter QM/MM MD simulations than in our previous investigation with standard QM/MM FEP calculations.[14] To test our method, we studied the MM→QM/MM contribution to the binding free energies for nine cyclic carboxylate ligands bound to the octa-acid host from the SAMPL4 competition. We found that the method performs well for eight out of the nine ligands. For one ligand, the convergence is slower and the problem was traced to too large a difference in the MM and QM/MM potentials which consequently required the QM/MM MD simulations to overcome a slow conformational change.

We found that it was not necessary to use QM/MM minimizations before running the MD simulations. However, we find that the aromatic ligands gain somewhat from removing the first 1–5 ps of the simulations as equilibration. Overlap measures were computed for

all ligands in the four Λ-windows and it was found that all of our results had appropriately overlapping energy distributions.

We finally note that with this approach, the efficiency is comparable to employing Jarzynski's equation for the non-equilibrium work, which has been employed in some studies.[44] In a future publication, we will compare the accuracy, precision and efficiency of these two approaches.

Acknowledgement

This work was supported by the Swedish Research Council (project 2014-5540) and the Knut and Alice Wallenberg Foundation (KAW 2013.022). C. S. thanks the Danish Council for Independent Research (the Sapere Aude program) for financial support (Grant no. DFF 4181-00370). 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

Supporting Information Available

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(2) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(3) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

(4) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107.

(5) Gilson, M.; Given, J.; Bush, B.; McCammon, J. *Biophys. J.* **1997**, *72*, 1047–1069.

(6) Hansen, N.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2014**, *10*, 2632–2647.

(7) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2009**, NA–NA.

(8) Ryde, U.; Söderhjelm, P. *Chem. Rev.* **2016**, *116*, 5520–5566.

(9) Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

(10) Ryde, U. In *Computational Approaches for Studying Enzyme Mechanism Part A*; Voth, G. A., Ed.; Methods in Enzymology; Elsevier, 2016; Vol. 577; Chapter 6, pp 119–158.

(11) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 9296–9297.

(12) Rathore, R. S.; Reddy, R. N.; Kondapi, A. K.; Reddanna, P.; Reddy, M. R. *Theor. Chem. Acc.* **2012**, *131*.

(13) Świderek, K.; Martí, S.; Moliner, V. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12614.

(14) Olsson, M. A.; Ryde, U. *J. Chem. Theory Comput.* **2017**, *13*, 2245–2253.

(15) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199–213.

(16) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94*.

(17) Rod, T. H.; Ryde, U. *J. Chem. Theory Comput.* **2005**, *1*, 1240–1251.

(18) Duarte, F.; Amrein, B. A.; Blaha-Nelson, D.; Kamerlin, S. C. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 954–965.

(19) Olsson, M. A.; Söderhjelm, P.; Ryde, U. *J. Comput. Chem.* **2016**, *37*, 1589–1600.

(20) Knig, G.; Boresch, S. *J. Comput. Chem.* **2011**, *32*, 1082–1090.

(21) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. *J. Chem. Theory Comput.* **2016**, *12*, 499–511.

(22) Beierlein, F. R.; Michel, J.; Essex, J. W. *J. Phys. Chem. B.* **2011**, *115*, 4911–4926.

(23) Mikulskis, P.; Cioloboc, D.; Andrejić, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Söderhjelm, P.; Ryde, U. *J. Comput. Aided Mol. Des.* **2014**, *28*, 375–400.

(24) Genheden, S.; Ryde, U.; Söderhjelm, P. *J. Comput. Chem.* **2015**, *36*, 2114–2124.

(25) Sampson, C.; Fox, T.; Tautermann, C. S.; Woods, C.; Skylaris, C.-K. *J. Phys. Chem. B.* **2015**, *119*, 7030–7040.

(26) Hummer, G.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 2004–2010.

(27) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J. Chem. Theory Comput.* **2006**, *2*, 452–461.

(28) Woods, C. J.; Shaw, K. E.; Mulholland, A. J. *J. Phys. Chem. B.* **2015**, *119*, 997–1001.

(29) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B.* **2011**, *115*, 7950–7962.

(30) Plotnikov, N. V.; Warshel, A. *J. Phys. Chem. B.* **2012**, *116*, 10342–10356.

(31) Lameira, J.; Kupchencko, I.; Warshel, A. *J. Phys. Chem. B.* **2016**, *120*, 2155–2164.

(32) Gibb, C. L. D.; Gibb, B. C. *J. Comput. Aided Mol. Des.* **2013**, *28*, 319–325.

(33) Muddana, H. S.; Fenley, A. T.; Mobley, D. L.; Gilson, M. K. *J. Comput. Aided Mol. Des.* **2014**, *28*, 305–317.

(34) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(35) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(36) Case, D. A.; Betz, R.; Cerutti, D. S.; Cheatham, T. E.; III,; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *University of California, San Francisco.* **2014**,

(37) Korth, M. *J. Chem. Theory Comput.* **2010**, *6*, 3808–3816.

(38) Wu, X.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381*, 512–518.

(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(40) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

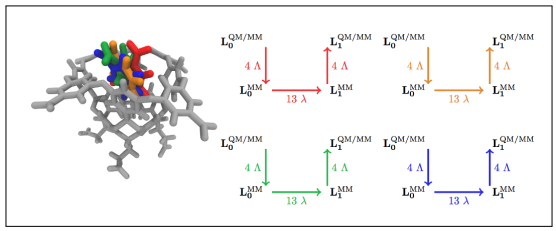
(41) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.

(42) Nicholls, A. *J. Comput. Aided Mol. Des.* **2016**, *30*, 103–126.

(43) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.

(44) Hudson, P. S.; Woodcock, H. L.; Boresch, S. *J. Phys. Chem. Lett.* **2015**, *6*, 4850–4856.

Graphical TOC Entry



Paper IV

Octav Caldararu, Martin A. Olsson, Majda Misini Ignjatović, Meiting Wang, Ulf Ryde, Manuscript, **2018**

Binding free energies in the SAMPL6 octa-acid host–guest challenge calculated with MM and QM methods

Octav Caldararu, Martin A. Olsson, Majda Misini Ignjatović, Meiting Wang and Ulf Ryde *

Department of Theoretical 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 222864

Key Words: ligand binding, free-energy perturbation, reference-potential with QM/MM sampling, semiempirical methods, density functional theory, entropy

ABSTRACT: We have estimated the free energy for the binding of eight carboxylic ligands to two variants of the octa-acid deep-cavity host in the SAMPL6 blind-test competition. We employed four different methods: Free-energy perturbation (FEP) for relative binding energies at the molecular mechanics (MM) and the combined quantum mechanical (QM) and MM (QM/MM) levels, the latter obtained with the reference-potential approach with QM/MM sampling for the MM→QM/MM FEP. The semiempirical PM6-DH+ method was employed for the ligand in these calculations. Moreover, binding free energies were estimated from QM/MM optimised structures, combined with COSMO-RS estimates of the solvation energy and thermostatistical corrections from MM frequencies. They were performed at the PM6-DH+ level of theory with the full host and guest molecule in the QM system (and also four water molecules in the geometry optimisations) and they were performed for 10–20 snapshots from molecular dynamics simulations of the complex. Finally, the structure with the lowest free energy was recalculated, using the dispersion-corrected density-functional theory method TPSS-D3, both for the structure and the energy. The two FEP approaches gave similar results (QM/MM slightly better for one of the two host), which were among the best five submissions in the challenge, with mean absolute deviations (MAD) of 2.4–5.2 kJ/mol and a correlation coefficient (R^2) of 0.77–0.93. This is the first time QM/MM approaches give results that are competitive to MM for the octa-acid host. The QM/MM-optimised structures gave somewhat worse results (MAD = 5–8 kJ/mol and R^2 = 0.4–0.7), but the results were improved compared to previous studies of this system with similar methods.

INTRODUCTION

Estimating the affinity between a small molecule and a biomacromolecule is important in many parts of chemistry, especially in drug design [1, 2]. Therefore, numerous computational methods have been developed with this aim [1], ranging from simple scoring approaches for ligand docking [3], via end-point approaches, like linear interaction energy [4] and MM/PBSA (molecular mechanics combined with Poisson–Boltzmann and surface area solvation) [5, 6], to strict approaches based on free-energy simulations [7, 8] with free energies calculated by exponential averaging (EA) [9], thermodynamic integration [10] or the Bennett acceptance ratio (BAR) approach [11].

The latter methods should in principle be limited only by the accuracy of the potential-energy function and the sampling of the phase space, although uncertainties in the nature of the simulated system (e.g. the protonation state of all involved molecules and residues) may also affect the results [7, 8]. To reduce the latter two types of problems, there have been quite some interest to study simpler systems, in particular the binding of small molecules to organic molecules of intermediate size (a few hundred atoms), i.e. host–guest systems [12, 13].

Most free-energy simulations are performed by empirical potentials in the form of molecular mechanics (MM) force fields. However, during the latest decades, there has been an increasing interest in employing quantum mechanical (QM) calculations to obtain more accurate energies [14]. Again, such calculations can be performed at many levels of approximation. Owing to the much larger computational cost of QM calculations, most such studies are based either on single-point calculations on structures obtained by MM sampling or on structures minimised by QM [14–16] (or rather by combined QM and MM calculations, QM/MM [17]). Only a few cases have involved sampling at the QM/MM level [18–22].

Most computational studies of ligand-binding affinities are performed on systems for which the experimental affinities are known. Of course, this introduce the risk that the results are biased towards the experimental data. Therefore, prospective studies, in which the experimental results are not known when the calculations are performed, provide a more unbiased view of the performance of various methods. In this regard, the SAMPL blind-test competitions have been invaluable to compare the true predictive value of various computational methods. Since, SAMPL3, it has involved host–guest systems [23] and since SAMPL4, the binding of ligands to the octa-acid deep-cavity host (OAH) [24, 25], developed by the Gibb group [26, 27].

In a series of publications, we have studied the binding of nine cyclic carboxylate guest molecules to the octa-acid host with computational method at both the MM and QM level [15, 21, 22, 28–30]. In the SAMPL4 competition, we used free-energy perturbation (FEP) to calculate the relative affinities of the nine guests at the MM level [15], which gave the best results in SAMPL4 [24], with a mean absolute deviation (MAD) of 3.6±0.2 kJ/mol and a correlation (R^2) to experimental data of 0.84±0.04. We also tried to improve the FEP results by performing QM calculations with density-functional theory (DFT) on snapshots from the MM simulations, using large QM systems involving ~310 atoms. However, the difference between the MM and QM potentials were so large that no converged results could be obtained. Therefore, the results were very poor with an uncertainty of 6–32 kJ/mol and MADs of 17–27 kJ/mol.

By using smaller QM systems and semiempirical calculations, we were later able to obtain converged results with a precision of 1 kJ/mol for all relative free energies, using 700 000 QM calculations for each ligand [28]. However, the results were still worse than the FEP results, with a MAD of 4.9±0.4 kJ/mol and a vanishing correlation. These results were obtained without any sampling at the QM/MM level, but in our next study such sampling was performed (with semiempirical PM6-DH+ calculations and only the ligand in the QM system) [21]. This gave even better results with a precision of 0.5–0.9 kJ/mol, a MAD of 4.7±0.2 and a R^2 correlation of 0.86±0.04. Recently, we have shown that similar results can be obtained with four times less computational effort using multiple short QM/MM simulations [22].

In the SAMPL4 study, we also tried to estimate octa-acid binding affinities with minimised QM structures, using the approach suggested by Grimme and coworkers [15]. We optimised the structures of the complexes with three different DFT approaches (in vacuum, in a continuum solvent and in a continuum solvent with four explicit water molecules). Then, binding free energies were calculated with a vacuum DFT calculation with large basis set and empirical dispersion corrections, combined by a COSMO-RS estimation of the solvation free energy and with thermostatistical corrections from a frequency calculation at the MM level. This approach gave absolute binding affinities of an intermediate accuracy with MADs of 7–14 kJ/mol and R^2 of 0.60–0.78. After removing systematic errors (the mean signed deviation), the MADs (called MADtr in the following) were 5–9 kJ/mol. Similar results were obtained also by Sure and Grimme on the same system [31]. An attempt to improve the energies by local coupled-cluster calculations gave much worse results with MAD, MADtr and R^2 of 37, 14 and 0.28 [15, 29].

In SAMPL5, we employed a similar approach to calculate binding affinities of six much more diverse guest molecules (with either a carboxylate or a tetramethylamine group) [30] to the octa-acid host and also to its tetra-endo-methyl variant (OAM) [32]. The calculations were improved by keeping the structures as symmetric as possible, reducing the charge and flexibility of the ligand and performing a restricted sampling of the complexes. Quite disappointing, the results were worse than for SAMPL4 with MADtr of 11–22 kJ/mol and R^2 below 0.30. The reason for this is probably the larger diversity of the ligands but also problems with some of the geometry optimisations. The results were not improved by employing DLPNO-CCSD(T) calculations [33] (MADtr of 16–20 kJ/mol and R^2 of 0–0.15. The best results in the SAMPL5 competition were obtained for free-energy simulations at the MM level, dragging the ligand out of the host [25].

In this paper, we study the binding of eight carboxylic ligands to both the OAH and OAM hosts with four different methods: FEP at the MM level, FEP at the PM6-DH+/MM level, as well as optimised structures at the PM6-DH+ and DFT levels of theory. For the latter, we used more extensive sampling at the MM level and QM/MM optimised structures with explicit solvent. We also re-examine the SAMPL4 and 5 test cases with the third method to show that it gives improved results. For the first time, we get a slight improvement of the MM FEP results by employing the QM/MM correction.

METHODS

Setup of the studied systems. We have considered the eight ligands of SAMPL6, G0–G7, as well as three aliphatic carboxylates with 5–8 carbon atoms (called A5–A8) and the MeBz ligand from SAMPL4 [34]. They are shown in Figure 1. For some test calculations, we used also all nine cyclic carboxylate ligands from SAMPL4 [34] and the four carboxylate ligands from SAMPL5 [35] (shown in Figures S1 and S2). The host–guest complexes for the calculations were built from the coordinates for the octa-acid host with the guest molecules from previous blind-prediction challenges [15, 30]. The guest molecules were prepared and modified using the Avogadro software [36] and the geometry of the guest molecules was optimised with the UFF force field [37]. The OAM was constructed by adding four methyl groups at the corresponding hydrogen positions on the upper rim of OAH.

Charges for the two host molecules have been reported before [15, 30]. Charges for the ligands were obtained with the same restrained electrostatic potential approach [38]: The molecules were optimized with the

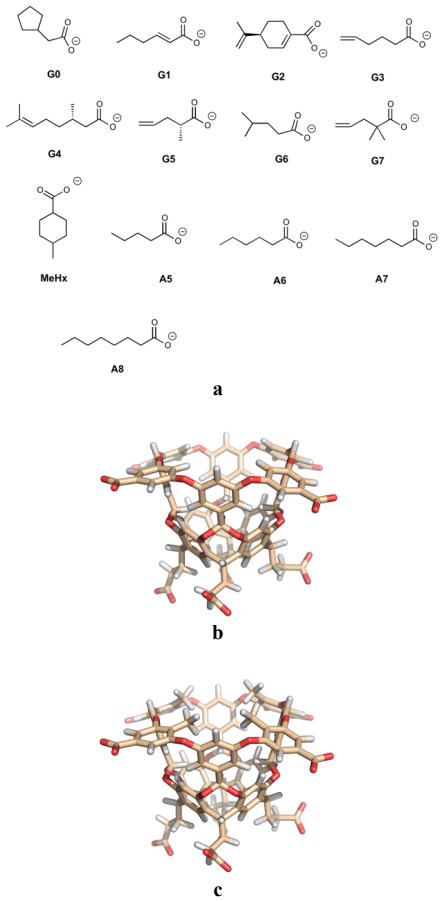


Figure 1. (a) Ligands involved in SAMPL6 challenge (G0–G7) and added to make the perturbations smaller and connected to experimental data (MeHx) and A5–A8. (b) The OAH and (c) OAM host molecules.

semiempirical AM1 method [39], followed by a single-point calculation at the Hartree–Fock/6-31G* level to obtain the electrostatic potentials, sampled with the Merz–Kollman scheme [40], but at a higher-than-default density (10 layers with 17 points per unit area, giving ~2 000 points per atom). These calculations were performed with the Gaussian 09 software [41]. The potentials were then used by antechamber to fit the charges. The charges and atom types of all ligands are given in Table S3 in the supplementary material.

A few parameters were missing in the force field and they were estimated with the Seminario approach [42]: The geometry of the ligands was optimised at TPSS/def2-SV(P) level [43, 44], followed by a frequency calculation using the aoforce module of the Turbomole software [45]. From the resulting Hessian matrix, parameters for the missing dihedrals were extracted with the Hess2FF program [46]. These parameters are given in Table S1 in the supplementary material.

Molecular dynamics simulation. All molecular dynamics (MD) simulations and FEP calculations were run with the AMBER 16 software suite [47]. Each host–guest complex was solvated in an octahedral box of water molecules extending at least 10 Å from the guest molecules using the tleap module, so that 1504–1513 water molecules were included in the simulations. All nine carboxylic groups on the host and guest molecules were assumed to be deprotonated because the binding affinities were measured at a pH of 11. Thus, the net charge of the host–guest complexes were –9. No counter ions were used in the simulations, as our previous studies have shown that they have a small effect on the calculated free energies [15]. Both the host and the guest molecules were treated with the general AMBER force field (GAFF) [48], whereas the TIP3P model was used for water molecules [49].

Each complex was minimised by 10 000 steps, followed by 20 ps constant-volume equilibration and 20 ps constant-pressure equilibration, all performed with heavy non-water atoms restrained towards the starting structure with a force constant of 209 kJ/mol/Å^2. Finally, the system was equilibrated for 2 ns without any restraints and with constant pressure, followed by 10 ns of production simulation, during which coordinates were saved every 5 or 10 ps. For each host–guest complex, 10 (OAH) or 20 (OAM) independent simulations were run, employing different TIP3P solvation boxes and different starting velocities [50]. Consequently, the total simulation time for each complex was 100 or 200 ns.

All bonds involving hydrogen atoms were constrained to the equilibrium value using the SHAKE algorithm [51], allowing for a time step of 2 ps. The temperature was kept constant at 300 K using Langevin dynamics [52], with a collision frequency of 2 ps^{-1}. The pressure was kept constant at 1 atm using a weak-coupling isotropic algorithm [53] with a relaxation time of 1 ps. Long-range electrostatics were handled by particle-mesh Ewald summation [54] with a fourth-order B spline interpolation and a tolerance of 10^{-5}. The cut-off radius for Lennard–Jones interactions was set to 8 Å.

Free-energy perturbations. The guest molecules were manually mapped for the FEP simulations as is shown in Figure 2, keeping the perturbations as small as possible. To this aim and also to connect the relative FEP calculations to experimental data [34, 55], we included also the A6–A8 and MeBz ligands. The FEP simulations were run with the pmemd module of AMBER 16 [37], using the dual-topology scheme with both ligands in the topology file. Each ligand transformation was divided into steps 13 steps, employing a linear transformation of the force-field potentials with the coupling parameter λ = 0.00, 0.05, 0.10, 0.20, …, 0.80, 0.90, 0.95 and 1.00. Electrostatic and van der Waals interactions were perturbed concomitantly, using soft-core potentials for both types of interactions [56, 57]. Soft-core potentials were used not only for atoms differing between the two ligands, but also for all atoms in the ring systems of ligands G0, G2 and MeHx to allow for larger differences in the

dynamics of the perturbed groups (atoms without soft-core potentials have identical coordinates in the perturbations).

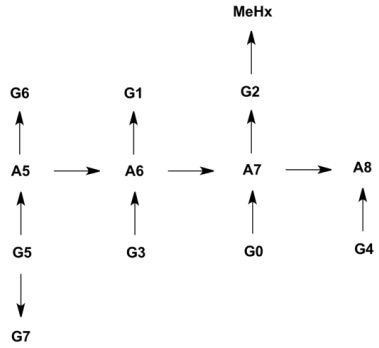


Figure 2. Ligand alchemical transformations studied with FEP.

For each λ value, 100 steps of minimisation were performed with the heavy atoms of the host and ligand restrained towards the starting structure with a force constant of 418 kJ/mol/Å^2. This was followed by 20 ps constant-volume equilibration with the same restraints and 1 ns constant-pressure equilibration without any restraints. Finally, a 2-ns production simulation was run (still with constant pressure), during which structures and energies were sampled every 2 ps.

Relative binding free energies between two ligands, L_0 and L_1 ($\Delta\Delta G_{bind}$), were estimated using a thermodynamic cycle that relates $\Delta\Delta G_{bind}$ to the free energy of alchemically transforming L_0 into L_1 when they were either bound to the host, $\Delta\Delta G_{bound}$, or were free in solution, $\Delta\Delta G_{free}$ [58, 59]:

$$\Delta\Delta G_{bind} = \Delta G_{bind}(L_1) - \Delta G_{bind}(L_0) = \Delta\Delta G_{bound} - \Delta\Delta G_{free} \quad (1)$$

$\Delta\Delta G_{bound}$ and $\Delta\Delta G_{free}$ were estimated by the multi-state Bennett acceptance-ratio (MBAR) method, using the pyMBAR software [60], including only statistically non-correlated energies in the calculations. All FEP calculations were repeated three times using different TIP3P solvation boxes and different starting velocities [50]. Reported free energies are the average over these three calculations, whereas the reported uncertainty is either the standard deviation over these three calculations divided by the square root of three or the square root of the variances of the three individual estimates divided by three, depending on which of the values was largest.

QM/MM FEP calculations. Relative QM/MM binding affinities between two ligands, L_0 and L_1, were estimated by the reference-potential method with QM/MM sampling (RPQS) [21, 22]. In this approach, the $\Delta\Delta G_{bind}$ free energies, calculated at the MM level, as described in the previous section, are corrected by a FEP calculation for each ligand in the methods space, from the MM potential to the QM/MM potential, as is shown by the thermodynamic cycle in Figure 3. This was done both for the ligand bound to the host and when free in solution. For each state (s = bound or free), the QM/MM corrected free energy was calculated from

$$\Delta\Delta G^{QM/MM}_{L_0\to L_1,s} = \Delta\Delta G^{MM}_{L_0\to L_1,s} - \Delta\Delta G^{MM\to QM/MM}_{L_0,s} + \Delta\Delta G^{MM\to QM/MM}_{L_1,s} \quad (2)$$

Finally, the net binding free energies were calculated from

$$\Delta\Delta G^{QM/MM}_{bind} = \Delta G^{QM/MM}_{L_0\to L_1,bound} - \Delta G^{QM/MM}_{L_0\to L_1,free} \quad (3)$$

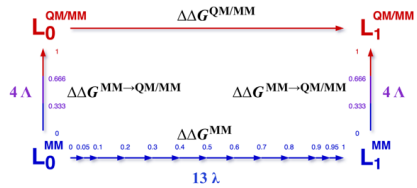


Figure 3. Thermodynamic cycle used for the RPQS calculations.

All MM → QM/MM FEP simulations were performed with the AMBER 16 software [47]. The calculations were performed for all host–guest systems involved in this study except for the MeHx ligand. In the QM/MM calculations, only the guest molecule was included in the QM region it was treated at the semiempirical PM6-DH+ level of theory [61–63]. The MM → QM/MM free energies were calculated based on the energy function $E(\Lambda) = (1 - \Lambda)E_{MM} + \Lambda E_{QM/MM}$, where E_{MM} is the MM energy, $E_{QM/MM}$ is the QM/MM energy and Λ is a coupling parameter going from 0 to 1. Based on our previous study of OAH with the SAMPL4 ligand [21], we performed calculations at four Λ values: 0.0, 0.333, 0.666, and 1.0. In some cases, additional Λ values were employed (0.166, 0.5 or 0.833), if the overlap with four Λ was unsatisfactory.

For each Λ value, we performed 100 steps of minimization with the heavy atoms of the host and guest molecules restrained towards the starting structure with a force constant of 418 kJ/mol/Å^2. This was followed by 20 ps constant-volume equilibration with the same restraints and 0.5 ns constant-pressure equilibration without any restraints. Finally, a 1 ns production simulation was run, during which structures and energies were sampled every 1 ps. The MD simulations were performed as described above, except that no bonds were constrained and the time step was 1 fs.

Absolute binding free energies from QM/MM minimised structures. Absolute binding free energies were calculated using the method suggested by Grimme [16, 64], in which the binding free energy is composed of three terms:

$$\Delta G_{tot} = \Delta E_{QM} + \Delta G_{solv} + \Delta G_{therm} \quad (4)$$

where ΔE_{QM} is a single-point vacuum QM energy, which also includes the dispersion energy, ΔG_{solv} is the

solvation free energy and ΔG_{therm} is a thermostatistical correction term. The binding affinity was obtained as the difference in this free energy between the complex, host and guest:

$$\Delta G_{\text{bind}} = \Delta G_{\text{tot}}(\text{complex}) - \Delta G_{\text{tot}}(\text{host}) - \Delta G_{\text{tot}}(\text{guest}) \quad (5)$$

Structures of the free host and guest molecules were taken from the structures of the complexes, without further optimisation (rigid binding free energies; some tests were performed to calculate structure relaxation energies, but they did not lead to any improvement). The calculations were performed at two levels of QM theory and based on two sets of structures. The two approaches will be called SQM and DFT in the following.

From each of the MD simulations of the host–guest complexes, the last snapshot was minimised at the PM6-DH+/MM level of theory [61, 62, 65] using the AMBER 16 software suite [47]. The quantum system consisted of the host and guest molecules, as well as four water molecules that formed hydrogen bonds with the guest (viz. the two closest water molecules to each of the carboxylate oxygen atoms). It had a net-charge of –9. The solvation box from the MD simulations was kept in all calculations. Minimisations were run for 2000 steps, with no periodicity (for technical reasons). This gave ten different host–guest structures for each guest bound to the OAH host and 20 different structures for the OAM host. The resulting structures were used directly for the SQM calculations.

The QM energy for the SQM structures (only isolated host and guest, with waters removed) was calculated as a PM6-DH+ single-point energy using the AMBER sqm program [47]. This method includes dispersion and hydrogen-bond corrections [61, 62, 65].

Solvation free energies in water solution were calculated with the conductor-like solvent model (COSMO) [66, 67] real-solvent (COSMO-RS) approach [68, 69] using the COSMOTHERM software [70]. These calculations were based on two single-point BP86 calculations [71, 72] with the TZVP basis set [73], one performed in a vacuum and the other in the COSMO solvent with an infinite dielectric constant. Owing to the extensive negative charge of the hosts, we had to use the undocumented ADEG option to force the program to accept that the solvation energy is very large.

Thermal corrections to the Gibbs free energy at 298 K and 1 atm pressure (ΔG_{therm}), including zero-point vibrational energy, entropy and enthalpy corrections, were calculated by an ideal-gas rigid-rotor harmonic-oscillator approach [74] from vibrational frequencies calculated at the MM level. To obtain more stable results, low-lying vibrational modes were treated by the free-rotor approximation, using the interpolation model suggested by Grimme with $\omega_0 = 100$ cm^{-1} [16]. The translational entropy was corrected by 7.99 kJ/mol for the change in the standard state from 1 atm to 1 M (used in the experiments).

For the SQM calculations, these energies, obtained according to Eqns. 4 and 5 were calculated for all 10 or 20 snapshots and the final absolute ΔG_{bind} energy was obtained by either taking the minimum value, the average value or the Boltzmann-weighted average value.

In the second (DFT) approach, the structure with the most favourable SQM ΔG_{bind} energy was selected and it was further optimised at the QM/MM level with the same QM system, but with the TPSS-D3/def2-SV(P) method [43, 44]. These calculations were performed with the ComQum program [75, 76], which is an interface between AMBER [47] and the QM software Turbomole software [45]. In these calculations, the MM system was kept fixed. All complexes converged within 150 geometry iterations.

For optimised structures, ΔE_{QM} was calculated with the TPSS functional and the def2-QZVP' basis set (the def2-QZVP basis set [44] with the *f*-type functions on hydrogen and the *g*-type functions on the other atoms deleted). The dispersion energy was included using the DFT-D3 approach [77] with Becke–Johnson damping [78] and third-order terms included. All DFT calculations were sped up by expanding the Coulomb interactions in auxiliary basis sets with the resolution-of-identity approximation (RI), using the corresponding auxiliary basis sets [79, 80]. The multipole-accelerated resolution-of-identity J approach was also employed [81]. All DFT calculations were performed using the Turbomole 7.1 or 7.2 software [45]. Finally, absolute ΔG_{bind} energies were obtained with Eqns. 4 and 5, using the same approach to get ΔG_{solv} and ΔG_{therm} as for the SQM structures. However, the final ΔG_{bind} was based on a single DFT structure.

Error estimates, quality and overlap measures. All reported uncertainties are standard errors of the mean (standard deviations divided by the square root of the number of samples). The uncertainty of the MBAR free energies calculated at each λ or Λ value was estimated by bootstrapping using the PYMBAR software [82] and the total uncertainty was obtained by error propagation.

The performance of the free-energy estimates was quantified by the mean signed deviation (MSD), the mean absolute deviation (MAD), the MAD after removal of the MSD (MADtr), the root-mean-square deviation (RMSD), the maximum error (Max), the correlation coefficient (R^2), the slope of the best correlation line and Kendall's rank correlation coefficient (τ) compared to experimental data. For relative affinities, τ was calculated only for the transformations that were explicitly studied, not for all combinations that can be formed from these transformations (this is marked by calling it τ_r). Moreover, it was also evaluated considering only differences (both experimental and calculated) that are statistically significant at the 90% level (τ_{90} and $\tau_{r,90}$ for absolute and relative affinities, respectively) [83]. Note that R^2 and the slope for relative affinities depend on the direction of the perturbation (i.e. whether $L_0 \rightarrow L_1$ or $L_1 \rightarrow L_0$ is considered, which is arbitrary). This was solved by considering both directions (both forward and backward) for all perturbations when these two measures were calculated.

The standard deviation of the quality measures was obtained by a simple simulation approach [84]. For

each transformation, 1000 Gaussian-distributed random numbers were generated with the mean and standard deviation equal to the MBAR and experimental results for that transformation. Then, the quality measures were calculated for each of these 1000 sets of simulated results and the standard error over the 1000 sets is reported as the uncertainty.

For all λ and Λ values of all perturbations, we have monitored five overlap measures, to ensure that the overlap of the studied distributions is satisfactory, viz. the Bhattacharyya coefficient Ω [85] the Wu and Kofke overlap measures of the energy probability distributions (K_{AB}) [86] and their bias metrics (Π) [86], the weight of the maximum term in the exponential average (w_{max}) [87] and the difference of the forward and backward exponential average estimate ($\Delta\Delta G_{EA}$) [88]. If $\Pi < 0$ or two of the following criteria were not fulfilled: $\Omega > 0.7$, $K_{AB} > 0.7$, $\Pi > 0.5$, $w_{max} < 0.2$, $\Delta\Delta G_{EA} < 4$ kJ/mol, additional λ or Λ values were included.

RESULT AND DISCUSSION

In this study, we have calculated the free energies for the binding of the eight ligands G0–G8 in Figure 1 to the normal (OAH) and methylated (OAM) deep-cavity octa-acid hosts (also shown in Figure 1) within the SAMPL6 blind-test competition. Thus, the experimental data were not known when the calculations were performed and was revealed only after the predictions were submitted. We employed four different methods and submitted four data sets: First, we performed standard relative FEP calculations at the MM level. Second, we performed QM/MM FEP calculations using the reference-potential approach with explicit QM/MM sampling (RPQS) [21] at the PM6-DH+ level of theory. Third, absolute binding free energies were estimated by semiempirical QM/MM optimisations on 10–20 snapshots from MD simulations and with PM6-DH+ energies supplemented by continuum solvation and thermostatistical corrections. Fourth, the best of the latter structures was reoptimised at the DFT level and energies were calculated with DFT and large basis sets. The results are described below in separate subsections.

FEP calculations at the MM level. We have calculated the relative binding free energies of the SAMPL6 ligands G0–G7 by FEP calculations at the MM level. As can be seen in Figure 1, the eight ligands contain a carboxylic group and five to ten carbon atoms. G0 and G2 involve a five- or six-membered ring and all except G0 and G6 have one or two double bonds. Ligands G2, G4 and G5 are chiral and they were used in the isomers shown in Figure 1 (since the host is achiral, the actual form should not matter for the binding affinities).

We developed a FEP scheme, shown in Figure 2, in which the eight ligands are connected, keeping the change as small as possible. This was partly accomplished by adding four extra ligands, which are the aliphatic carboxylates with five to eight carbon atoms, A5–A8. Thereby, the perturbations are restricted to the introduction of a double bound, the conversion of a H atom to a methyl group, the closure of a ring, or in one case (G2), formation of a cyclohexene ring by the addition of two carbon atoms. The aliphatic ligands were employed also because experimental binding affinities are available for A6 and A8 to OAH [55], giving us the opportunity to convert the relative energies to absolute affinities. For the same reason, the MeHx ligand from SAMPL4 (shown in Figure 1a) was also added and connected to G2. To connect the calculations of OAH and OAM, and to obtain absolute affinities for the OAM ligands, we also converted OAH to OAM with and without the A6 ligand bound.

The calculated relative affinities are listed in Table 1 (free energies calculated with MBAR). It can be seen that the precision of most $\Delta\Delta G_{bind}$ estimates is low, 0.1–0.4 kJ/mol, owing to the use of three independent FEP calculations. This reflects that the three estimates give similar results, with a variation of up to 1.0 kJ/mol for OAH and 1.6 kJ/mol for OAM. However, for two perturbations with both hosts, G2 → MeHx and G4→A8, the variation is much larger (9–20 kJ/mol) and therefore the precision is much worse, 1.5–2.6 kJ/mol even if we employed six independent simulations for these four perturbations.

Besides the G5→G7 perturbation, the results in Table 1 are not directly comparable to the experimental data, because they involve the A5–A8 and MeHx ligands that are not involved in the SAMPL6 measurements. We have used two different approaches to solve this problem. For the submitted data, we employed previously published experimental data for A6, A8 and MeHx in OAH [34, 55] to calculate absolute affinities for all ligands. This is a bit risky, because $\Delta\Delta G_{bind}$ measured in different studies (at slightly different conditions) vary somewhat. For example, the experimental ΔG_{bind} of Hx to OAH, involved in SAMPL4 vary between 21.1 and 23.5 kJ/mol in two articles by the same group [34, 89] and the results for A6, A8 and A10 vary by 1.5–2.8 kJ/mol [55, 89] (we employed the newer data in this article).

Our initial calculations along these lines showed that the calculated data were somewhat problematic: As can be seen in Figure 2, the A6 and A8 ligands are connected by two perturbations, A6 → A7 → A8. However, the initial result for these perturbations was quite poor, –11.3±0.3 kJ/mol, compared to the difference in the experimental $\Delta\Delta G_{bind}$ for A6 and A8, –4.9 or –6.2 kJ/mol in the two experimental studies [55, 89]. We therefore rerun these two perturbations with the whole ligand included in the perturbed group (instead of only the differing atoms). For A7 → A8, this did not change the results significantly, as can be seen in Table 1 (entry "entire ligand"). However, for A6 → A7, the result changed by 4 kJ/mol, bringing the A6 → A8 estimate closer to experiments, –7.8±1.2 kJ/mol. Unfortunately, we did not have time to rerun all the other perturbations with the whole ligand in the soft-core group, but we used the latter results for the A6 → A7 perturbation and also corrected the corresponding results for OAM with the difference between the two A6 → A7 perturbations for OAH.

Absolute affinities calculated this way are shown

in Table 2, together with the reference ligands (from which the experimental data was taken, because there are several possibilities) and the experimental data (revealed after submission our results). It can be seen that the agreement is rather good: As can be seen in Figure 4a, the 16 predictions have errors of 1.9–9.7 kJ/mol. Consequently, the MAD is quite high, 5.6±0.3 and 6.2±0.3 kJ/mol for the two hosts. For most of the ligands, the predicted affinities are less negative than the experimental ones – the MSD is 5.0±0.3 and 2.0±0.4 kJ/mol for the two hosts. However, for G4 in both hosts and G2 in OAM, the opposite is true. If the systematic error is removed, the MAD is improved significantly. For OAH, the MADtr is good, 2.6±0.3 kJ/mol, whereas it is worse for OAM, 5.2±0.4 kJ/mol. The reason for this is probably that the experimental data employed to calculate the absolute affinities were all for OAH, so the results for OAM involve more perturbations and therefore the possibility of accumulation of errors.

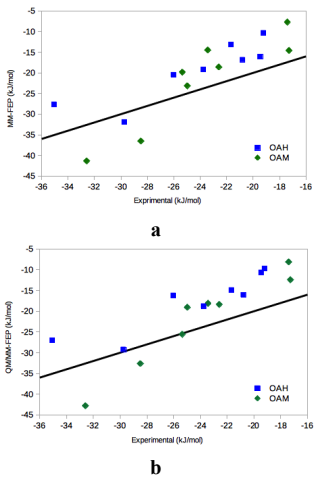


Figure 4. Comparison of the experimental and calculated absolute affinities obtained with the (a) MM-FEP and (b) QM/MM-FEP methods. The line shows the perfect correlation.

On the other hand, the correlation between the experimental and calculated results is better for OAM (R^2 = 0.85±0.02) than for OAH (0.77±0.05), although the difference is not fully significant. The same applies also for τ_{90}, which is 0.84±0.02 and 0.79±0.02 for OAM and OAH, respectively.

Alternatively, we instead considered only relative affinities. These were obtained by combining two or three perturbations so that they go only between the G0–G7 ligands. This can be done in a few different ways and one connected and consistent set of seven relative energies are shown in Table 3. It can be seen that the results are quite similar to those of the absolute affinities. The errors vary between 0.4 and 7.3 kJ/mol, except for the G0 → G2 difference in OAM, for which the error is as much as 13.6 kJ/mol (the calculated result overestimates the true difference, but with the correct sign). Consequently, the MAD is larger for OAM (5.1±0.2 kJ/mol) than for OAH (3.1±0.2 kJ/mol). R^2 is also better for OAH (0.87±0.02, compared to 0.61±0.04). On the other hand, $\tau_{r,90}$ is perfect for OAM (all statistically significant differences have the correct sign), whereas it is 0.71 for OAH (one difference has the incorrect sign). The single perturbation that involves only SAMPL6 ligands (G5 → G7) gives errors of the same size as the combined perturbations (3–7 kJ/mol), indicating that the results are not biased by poor performance of the added A5–A8 and MeHx ligands.

FEP calculations at the QM/MM level. Next, we used the RPQS approach to calculate all the relative binding affinities at the QM/MM level. For this, we performed MM → QM/MM FEP calculations for all G0–G7 and A5–A8 ligands both when bound to the host and free in solution (cf. Figure 3). The results are shown in Table 4. The individual $MM \rightarrow QM/MM$ free energies calculated when the ligand is bound to the host ($\Delta G_{L,bound}^{MM \rightarrow QM/MM}$) or free in solution ($\Delta G_{L,free}^{MM \rightarrow QM/MM}$), ranged from –507 to –691, except for G4 (around –260 kJ/mol) and G7 (around –962 kJ/mol). However, for each ligand, the values in the host and in solution were of a similar size, and the resulting $MM \rightarrow QM/MM$ correction to ΔG_{bind} ($\Delta\Delta G_{bind,L}^{MM \rightarrow QM/MM}$, shown in Table 4) ranged between –8.8 and +5.7 kJ/mol.

The standard errors are between 0.2 and 0.4 kJ/mol, except for G1 and G2 bound to OAM, for which they were 0.9 and 1.2 kJ/mol. For G0–G7, we run duplicate calculations and for these two ligands, the results differed by 1.9 and 2.4 kJ/mol, whereas for the other ligands, they agreed within 0.5 kJ/mol. In fact, the large variation came from the $\Delta G_{L,bound}^{MM \rightarrow QM/MM}$ term, which varied by 2.6 and 3.2 kJ/mol for these two ligands, but less than 0.2 kJ/mol for the other ligands. For the $\Delta G_{L,free}^{MM \rightarrow QM/MM}$ term, for which we have two or three samples of each, the variation was 0.1–0.8 kJ/mol, except for G2 and G6 (1.2 and 2.0 kJ/mol).

We used five overlap measures (described in the Method section) to check that the calculated MM → QM corrections are reliable. Based on these, we added intermediate Λ values for some of the ligands, as is shown in the last two columns of Table 4.

Next, the $\Delta\Delta G_{bind,L}^{MM \rightarrow QM/MM}$ corrections in Table 4 were combined with the results of the FEP calculations at the MM level ($\Delta\Delta G_{bind,L_0 \rightarrow L_1}^{MM}$ in Table 1) to get the final QM/MM relative binding free energies ($\Delta\Delta G_{L_0 \rightarrow L_1}^{QM/MM}$). These results are also included in Table 1. It can be seen that most MM → QM/MM corrections are rather small, 0.1–3.7 kJ/mol (average 2.2 kJ/mol). However, for the G1 → A6 and G2 → A7 perturbations they are –5.4 to –10.7 kJ/mol.

These relative energies were then recalculated to absolute affinities in the same way as for the FEP results at the MM level. These results, are shown in Table 2 and in Figure 4b. They differ from the MM results by 2 kJ/mol on average and a maximum difference of 5.7 kJ/mol. For OAH, the results are

consistently less negative than the experimental results, by 5.0–9.5 kJ/mol for all ligands except G4 (0.6 kJ/mol; MSD = 6.7±0.3 kJ/mol). Therefore, the MAD is rather high, 6.7±0.3 kJ/mol, but the MADtr is excellent, 2.4±0.4 kJ/mol. For OAM, the deviation is less systematic and more varying with a MSD of 1.9±0.4 kJ/mol, MAD = 5.5±0.4 kJ/mol, MADtr = 5.0±0.5 kJ/mol and a maximum error of 10.2 kJ/mol for G4. However, the correlation is better for OAM (R^2 = 0.93±0.02, compared to 0.81±0.04 for OAH) and τ_{90} is perfect for OAM, but 0.84±0.02 for OAH. Compared to the MM-FEP results, the performance for OAH is similar (MAD, MSD and Max are worse, but MADtr, R^2 and τ_{90} are better). However, for OAM, the QM/MM-FEP results are clearly better for all quality measures, except for the maximum error.

We also made the corresponding analysis for the relative energies in Table 3. The results are similar to those obtained for the absolute energies: The MAD is lower for OAH than for OAM (3.9±0.4 compared to 4.9±0.4 kJ/mol). However, the correlation coefficient (R^2) is better for OAM, 0.73±0.04, compared to 0.56±0.08. $\tau_{r,90}$ is perfect for both hosts. Compared to the MM-FEP results, the two methods have a similar performance for OAH (MAD, R^2 and Max are better for MM-FEP, MSD and $\tau_{r,90}$ is better for QM/MM-FEP), but QM/MM-FEP is better (or equal) for OAM for all quality measures.

Absolute binding affinities from minimised semi-empirical structures. Next, we tried to calculate absolute binding affinities for all the SAMPL6 host–guest complexes with QM-optimised structures, using a variation of an approach developed by Grimme [16, 64]. In the SAMPL5 study [30], we noticed that vacuum optimisations led to structures that had the guest carboxylate groups too much buried inside the host. This could only partly be remedied by using an implicit solvent method, such as COSMO [66, 67] or by including four explicit water molecules in the calculations. Therefore, in this study, we decided to base the calculations on snapshots from a long MD simulation of the complex. For each snapshot, the structure was minimised at the QM/MM level of theory, keeping all water molecules in the MM system and including the host, guest and four water molecules (that form hydrogen bonds with the carboxylate group of the guest) in the QM system. To start with, we performed 100 or 200 ns MD simulations for each host–guest complex and extracted 10 or 20 snapshots from these.

To make the calculations rapid, allowing for calibration also on the SAMPL4 and SAMPL5 structures, we chose to employ the semiempirical dispersion- and hydrogen-bond-corrected PM6-DH+ method for the QM calculations. This reduced the computational effort to 3–5 hours (single-core) for the QM/MM minimisations, compared to 2–4 weeks for the previous DFT optimisations. This could be in principle further sped up by using parallel calculations or by keeping the MM system fixed or restrained during the minimisation. After minimisation, single-point PM6-DH+ energies were calculated for the isolated host–guest complex and these energies were combined with COSMO-RS solvation energies and thermostatistical corrections from a MM frequency calculation, according to Eqn. 4. The PM6-DH+ energy and MM frequency calculations took only some tens of seconds to complete, leaving only the COSMO-RS solvation energy calculations as the computational bottleneck, as these can take as much as 1 day to converge (besides the initial MD simulations, which take about 5 hours per 10 ns on one GPU).

We started by testing the protocol on the nine cyclic carboxylates binding to OAH in the SAMPL4 competition [24] (Figure S1), the four carboxylic ligands binding to OAH and OAM in SAMPL5 [25] (S5-G1, S5-G2, S5-G4 and S5-G6, shown in Figure S2; we omitted the two positively charged ligands as all SAMPL6 ligands have a single negative charge), as well as A6**,** A8 and A10 binding to OAH [55].

As described above, the binding energies were obtained from 10–20 snapshots from the MD simulations. Therefore, we need to decide how these binding energies should be combined to single final estimate. To this end, we compared three different approaches: the averaged energy, the minimum energy and the Boltzmann-weighted averaged energy. The results are presented in Tables 5 and 6 and are shown in Figure 5. It can be seen that minimum and Boltzmann-averaged energies give similar results (the former are slightly better for SAMPL4, whereas the opposite is true for SAMPL5). However, in both cases, the averaged energies give much higher MAD, MSD and maximum errors than the other two approaches. However, when the systematic underestimation of the binding affinities (9–11 kJ/mol) has been removed, they actually give the lowest MADtr. Moreover, the slope, as well as R^2 and τ_{90} are always best for the averaged energies. Theoretically, Boltzmann averaging is the preferred approach, it gave the best results in SAMPL5 [32] and it is also the approach we used for the submitted energies. The better performance of the plain averages may indicate that the sampling was incomplete. The averaged energies have the advantage of giving an uncertainty. It is quite high for all ligands, 1–6 kJ/mol, again showing that much more snapshots are needed to reach reliable results.

We have previously studied the same systems with minimised QM structures, but using more expensive DFT-D3 methods. Quite amazingly, the new results are better: For SAMPL4, the MADtr are 3.4–5.2 kJ/mol, which are better than the previous DFT-D3 results, 4.6–8.6 kJ/mol. However, R^2 is lower, 0.31–0.65, compared to 0.60–0.78, whereas τ_{90} is worse for the minimum and Boltzmann-averaged energies (0.33–0.38, compared to 0.71–0.77), but that of the averaged energies is best, 0.92. For SAMPL5, all the new results are much better than the old DFT-D3 results: MADtr = 5.1–5.8 kJ/mol, compared to 11–21 kJ/mol, R^2 = 0.38–0.45, compared to 0–0.30 (and in many cases negative correlation), and τ_{90} = 0.41–0.65, compared to –0.33 to 0.33. However, it should be remembered that we did not include in this study the two ligands with trimethylamine groups, which gave problems in the previous study. Still, we believe that the present approach involves at least two advantages. First, the

geometries were optimised in water, including four water molecules in the quantum system, which resulted in a lower repulsion of the negative carboxylate groups and a more realistic binding pose of the guests. Secondly, we used vibrational frequencies from a MM method instead of the HF-3c method, which gave large systematic errors in the previous study: The current method always gave negative binding free energies, while DFT-D3 with HF-3C frequencies gave many positive absolute binding affinities.

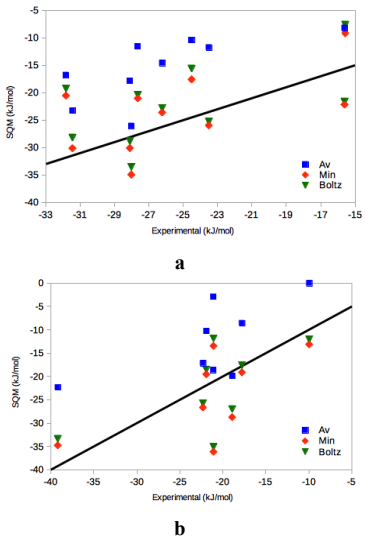


Figure 5. Comparison of the experimental and calculated absolute affinities obtained with the SQM method and three different ways to combine the ten energies from different snapshots, plain average (Av), the minimum energy (Min) or the Boltzmann average (Boltz) for the (a) SAMPL4 and (b) SAMPL5 ligands. The line shows the perfect correlation.

For the SAMPL6 ligands, the thermostatistical corrections from MM vibrational frequencies are quite stable, ranging only from 35 to 43 kJ/mol for all host–guest systems, except for about half of the OAM complexes with G2 and G7, for which they are 81–87 kJ/mol. The latter structures are all characterised with the ligand binding a bit deeper into the host. The PM6 energies are large and positive for all structures (748–986 kJ/mol), reflecting the large electrostatic repulsions between the carboxylate groups of the guests and the large negative charge (–8) of the hosts. However, they are always compensated by the COSMO-RS solvation energies, which are always negative (–826 to –1020 kJ/mol). Summing up these three terms gives the absolute binding free energy, which is always negative.

The results for the SQM approach using Boltzmann-weighted energies are collected in Table 7 and shown in Figure 6a. It can be seen that they are quite similar to the results obtained for the SAMPL4 and SAMPL5 tests, with errors of up to 14 kJ/mol for OAH and 24 kJ/mol for OAM. For OAH, the MAD is 6.6 kJ/mol and it is not much improved if the rather small systematic error (MSD = 1.2 kJ/mol) is removed (MADtr = 6.3 kJ/mol). The correlation is poor (R^2 = 0.18), but τ is slightly better, 0.36. For the OAM host, the MAD is significantly worse, 11.7 kJ/mol. However, this is caused by a systematic overestimation of all binding energies, except that of G3, giving a MSD of –10.7 kJ/mol. If this is removed, MADtr is only slightly higher than for OAH, 7.0 kJ/mol. Consequently, R^2 and τ_{90} are similar, 0.14 and 0.41, respectively.

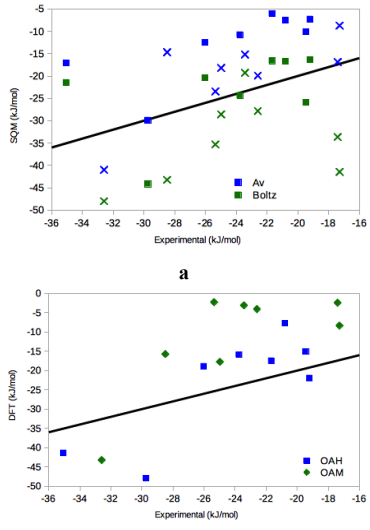


Figure 6. Comparison of the experimental and calculated absolute affinities obtained with the (a) SQM and (b) DFT methods, the former with two different ways to combine the 10–20 energies from different snapshots, plain average (Av) or the Boltzmann average (Boltz) for the SAMPL6 ligands. The line shows the perfect correlation. In (a), OAH energies are shown with squares and OAM energies with crosses.

Interestingly, the PM6-DH+ calculations suggest that all ligands bind stronger to the methylated OAM host than to the OAH host (by 3–22 kJ/mol). This is in contrast to the experimental data, which shows a much smaller difference between the two hosts (2–9 kJ/mol) and that ligands G2, G5 and G7 actually bind stronger to OAH. These are the three hosts with substituents at the C2 atom. Apparently, the PM6-DH+ energy function fails to model properly the competition between exchange repulsion and dispersion attraction between these substituents and the methyl groups on the host. In fact, PM6-DH+ instead predicts that these three ligands bind 17–22 kJ/mol better to OAM, which is more than for any of the other ligands. This gives rise to the largest errors for this approach with the OAM host (15–24 kJ/mol too strong binding). The binding of G2 is too weak to OAH, whereas G4 is predicted to bind too strongly to both hosts by 14–15 kJ/mol.

In fact, most results would have been improved if

we had selected to submit the results of the pure average instead. For OAH, they give a MADtr of 3.6±0.9 kJ/mol, a correlation of 0.51±0.16 and a τ_{90} of 0.83±0.07. This improvement comes mainly from a more favourable binding of G2 (compared to the other ligands), whereas G4 still shows a too strong binding. For OAM, MADtr is larger, 5.1±0.9 kJ/mol, whereas R^2 and τ_{90} are better, 0.55±0.15 and 0.86±0.06.

Absolute binding affinities from minimised DFT structures. Finally, we tried to improve the absolute binding affinities by using DFT calculations both in the geometry optimisations and in the energy estimates. Thus, we selected the minimum-energy snapshot according to SQM calculations and performed DFT/MM optimisation with the surrounding water included as a fixed MM system. We then calculated energies of the resulting structures in a similar way, using thermostatistical corrections from MM vibrational frequencies and COSMO-RS solvation energies, but with TPSS-d3/def2-QZVP' energies instead of the PM6-DH+ energies.

The DFT results are also included in Table 7 and they are shown Figure 6b. Unfortunately, they are considerably worse than the semiempirical QM results. In particular, most of the binding free energies were positive, mainly owing to a more positive thermostatistical correction (69–86 kJ/mol, again reflecting that the ligand is deeper bound in the DFT structures than in the PM6 structures). The solvation energies are of a similar magnitude in the two sets of calculations, whereas the QM energies are somewhat more positive for the PM6 calculations (by 11 kJ/mol on average). Therefore, the raw DFT energies give large MADs (22 kJ/mol for the OA host and 33 kJ/mol for the OAM host). We recognized this problem early and therefore submitted DFT energies with a constant offset, taken as the difference between the average PM6 and DFT results for the two hosts (–20.9 kJ/mol). This brought down the MAD to 8 and 15 kJ/mol for the two hosts, but of course did not affect most of the other quality measures. With the DFT calculations, the performance is similar for the two hosts, with MADtr of 7.5–7.7 kJ/mol, R^2 = 0.57–0.66 and τ_{90} = 0.33–0.43. Thus, the MADtr and τ_{90} are worse than for the SQM calculations, whereas R^2 is actually slightly better.

Comparison with other submissions. There were 43 submissions for the SAMPL6 octa-acid challenge from eight research groups (and one additional submission only for OAM). They were ranked for five quality measures, MAD, RMSD, MSE, R^2 and slope. Based on these, two submissions, employing potential-of-mean force umbrella sampling simulations (i.e. dragging the ligand out of the host) gave the best results with the lowest MAD and RMSD for both hosts (MAD = 1.7–2.1 kJ/mol for OAH and 4.3–4.7 kJ/mol for OAM), 3–8th best MSE, 7–10th best slope and 2–8th best R^2 (0.88–0.93), except for one of them on OAM (24th; 0.27). However, the same group submitted 27 different submissions with apparently the same method (no details are currently available), so it is unclear why these submissions are so good, whereas one of their submissions was the second worst in the competition.

The third best method seems to be a FEP study of relative free energies by the Michel group, employing GAFF with AM1-BCC charges, TIP3P water and no counter ions, i.e. similar to our MM-FEP approach, besides the charges. However, they obtained a better MAD for OAH (3.3, compared to 5.7 kJ/mol) and better MSD for both systems, whereas our MAD is better for OAM (5.5 compared to 7.2 kJ/mol). The same group submitted five data sets with similar methods (the differences are currently unclear, except that some involved counter ions). Two gave good results (third and sixth best), one intermediate (24th) and two poor (35–37th). Calculations without counter ions were always best.

Our MM-FEP and QM/MM-FEP gave similar results and were the fourth and fifth best among the submissions. They gave the 4–12th best results for all quality measures for OAM, whereas the performance on OAH was more mediocre (10–19th). However, our OAH results seems to somewhat affected by our use of experimental data to obtain absolute affinities: The MADtr for OAH for both MM- and QM/MM-FEP are very low, 2.4–2.6 kJ/mol, i.e. approaching the performance of the two best methods and clearly better than the third. The DFT and SQM results ranked in the lower middle, around positions 26–31, although the MSE was among the six best for OAH. It is likely that the performance would be improved if relative quality measures, like MADtr, were considered. However, it is quite satisfying that for the first time, QM approaches, like QM/MM-FEP come within the best five submissions. Moreover, both SQM and DFT gave decent results, of a quality similar to many of the MM FEP results, e.g. a submission employing FEP with the polarisable AMOEBA force field [90]. In particular, they are appreciably better than the other purely QM submission, employing B3PW91 calculations with complete basis sets and a SMD continuum solvent [91], which rank among the three worst results.

CONCLUSIONS

We have studied the binding of eight ligands to two variants of the octa-acid deep-cavity host in the SAMPL6 blind-test competition. We have employed four different approaches, three of which are based on QM methods. First, we performed standard relative FEP calculations at the MM level with free energies calculated with MBAR and employing the GAFF+TIP3P force fields and RESP charges. Second, we used the reference-potential approach with explicit QM/MM sampling to obtain relative FEP free energies at the semiempirical PM6-DH+/MM level of theory. Third, we employed the same semiempirical QM method to obtain QM/MM optimised structures, for which free energies were calculated by combining the PM6-DH+ interaction energies with COSMO-RS solvation free energies and thermostatistical corrections calculated at the MM level. We employed 10–20 structures taken from a MD simulation of the host–guest complexes. Finally, we reoptimised the best structures from the previous approach with the TPSS-D3/MM method and calculated QM energies with a

large basis set, which were then combined with COSMO-RS and thermostatistical corrections.

The MM- and QM/MM-FEP methods gave excellent results for OAH, with MADtr of 2.4–2.6 kJ/mol and R^2 of 0.77–0.81. For OAM, the MADtr is somewhat larger, 5.0–5.2 kJ/mol, but the R^2 is better, 0.85–0.93. For the former, the two approaches gave similar results, whereas for OAM, QM/MM-FEP was clearly better. These results were among the five best submissions to SAMPL6.

The SQM and DFT results were somewhat worse, especially for OAH; MADtr = 4–8 kJ/mol and R^2 = 0.55–0.66. Unfortunately, we selected to submit SQM results based on Boltzmann-averaged, rather than plain averaged energies, which gave somewhat worse results, MADtr = 6–7 kJ/mol and R^2 = 0.14–0.18. However, these methods gave similar results as our previous calculations with DFT-optimised structures in SAMPL4 and much better results for SAMPL5. Compared to the other submissions, these results were mediocre, but still comparable to many approaches employing MM-FEP methods. In particular, the were much better than another approach employing QM structures and energies.

The present results are satisfying because for the first time we are able to improve MM-FEP results for the octa-acid host with QM/MM methods and these results are among the best five submissions. These results were obtained with the simple and cheap semiempirical PM6-DH+ method, showing that it is much more important to perform a proper sampling and converge the MM→QM/MM FEP than employing a more accurate QM method. However, now when we have a working QM/MM-FEP method, the next challenge will be to extend it to more accurate QM methods and larger QM systems.

For the QM-minimised structures, we have shown that the results are improved by employing QM/MM-optimised structures, rather than structures optimised in vacuum or in a continuum solvent. This also made the calculations significantly faster. However, there are still several problems to solve with this approach. In particular, there seems to be a problem with absolute free energies, related to the entropy term, which vary by 10–40 kJ/mol, depending on what method is used for the geometries and the frequencies. In particular, we observe that the simple PM6-DH+ method gives better results that the inherently more accurate TPSS-D3 approach. Moreover, much more sampling seems to be needed before the results are stable and reliable. With 10–20 snapshots, plain averages gave better results (but with large uncertainties, 1–5 kJ/mol. Finally, improved methods to estimate the strain energies of the host and the guest in the complexes are needed.

Still it is very satisfying that QM-based methods start to have some impact also on calculated binding affinities for host–guest systems.

ACKOWLEDGEMENT

This investigation has been supported by grants from the Swedish research council (project 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. Gohlke H, Klebe G (2002) Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angew Chemie - Int Ed 41:2644–2676.
2. Jorgensen WL (2009) Effcient Drug Lead Discovery and Optimization. Acc Chem Res 42:724–733.
3. Kontoyianni M, Madhav P, Seibel ES (2008) Theoretical and Practical Considerations in Virtual Screening: A Beaten Field? Curr Med Chem 15:107–116.
4. Åqvist J, Luzhkov VB, Brandsdal BO (2002) Ligand Binding Affinities from MD Simulations. Acc Chem Res 35:358–365.
5. Kollman PA, Massova I, Reyes CM, et al (2000) Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc Chem Res 33:889–897.
6. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 441:1–13.
7. Wereszczynski J, McCammon JA (2012) Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. Q Rev Biophys 45:1–25.6
8. Hansen N, Van Gunsteren WF (2014) Practical aspects of free-energy calculations: A review. J Chem Theory Comput 10:2632–2647.
9. Zwanzig RW (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. J Chem Phys 22:1420–1426.
10. Kirkwood JG (1935) Statistical Mechanics of Fluid Mixtures. J Chem Phys 3:300.
11. Bennett CH (1976) MC method for FEP. Argues that it is always better to sample both ensembles than only one. Discusses overlap and efficiency. J Comput Phys 22:245–268.
12. Jensen JH (2015) Predicting accurate absolute binding energies in aqueous solution: thermodynamic considerations for electronic structure methods. Phys Chem Chem Phys 17:12441–51.
13. Moghaddam S, Inoue Y, Gilson MK (2009) Host–Guest Complexes with Protein–Ligand-like Affinities : Computational Analysis and Design Host - Guest Complexes with Protein - Ligand-like Affinities : 4012–4021.
14. Ryde U, Söderhjelm P (2016) Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. Chem Rev 116:5520–5566.
15. Mikulskis P, Cioloboc D, Andrejić M, et al (2014) Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host-guest binding energies. J Comput Aided Mol Des 28:375–400.
16. Grimme S (2012) Supramolecular binding thermodynamics by dispersion-corrected density functional theory. Chem - A Eur J 18:9955–9964.
17. Ryde U (2016) QM / MM Calculations on Proteins. Methods Enzymol 577:119–158.
18. Reddy MR, Erion MD (2007) Relative binding affinities of fructose-1,6-bisphosphatase inhibitors calculated using a quantum mechanics-based free energy perturbation method. J Am Chem Soc 129:9296–9297.
19. Rathore RS, Reddy RN, Kondapi AK, et al (2012) Use of quantum mechanics/molecular mechanics-based FEP method for calculating relative binding affinities of FBPase inhibitors for type-2 diabetes. Theor Chem Acc 131:1096; 10 pages.
20. Świderek K, Martí S, Moliner V (2012) Theoretical studies of HIV-1 reverse transcriptase inhibition. Phys Chem Chem Phys 14:12614–24.

21. Olsson MA, Ryde U (2017) Comparison of QM/MM Methods To Obtain Ligand-Binding Free Energies. J Chem Theory Comput 13:2245–2253.
22. Steinmann C, Olsson MA, Ryde U (2018) Relative ligand-binding free energies calculated from multiple short QM/MM MD simulations. J Chem Theory Comput submitted.
23. Muddana HS, Varnado CD, Bielawski CW, et al (2012) Blind prediction of host-guest binding affinities: A new SAMPL3 challenge. J Comput Aided Mol Des 26:475–487.
24. Muddana HS, Fenley AT, Mobley DL, Gilson MK (2014) The SAMPL4 host-guest blind prediction challenge: An overview. J Comput Aided Mol Des 28:305–317.
25. Yin J, Henriksen NM, Slochower DR, et al (2017) Overview of the SAMPL5 host – guest challenge : Are we doing better ? J Comput Aided Mol Des 31:1–19.
26. Xi H, L. D. Gibb C (1998) Deep-cavity cavitands: synthesis and solid state structure of host molecules possessing large bowl-shaped cavities. Chem Commun 1743–1744.
27. Liu S, Gibb BC (2008) High-definition self-assemblies driven by the hydrophobic effect: synthesis and properties of a supramolecular nanocapsule. Chem Commun (Camb) 7345:3709–16.
28. Olsson MA, Söderhjelm P, Ryde U (2016) Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level. J Comput Chem 37:1589–1600.
29. Andrejic M, Ryde U, Mata RA, Söderhjelm P (2014) Coupled-Cluster Interaction Energies for 200-Atom Host-Guest Systems. ChemPhysChem 15:3270–3281.
30. Caldararu O, Olsson MA, Riplinger C, et al (2017) Binding free energies in the SAMPL5 octa-acid host–guest challenge calculated with DFT-D3 and CCSD(T). J Comput Aided Mol Des 31:
31. Sure R, Antony J, Grimme S (2014) Blind prediction of binding affinities for charged supramolecular host-guest systems: Achievements and shortcomings of DFT-D3. J Phys Chem B 118:3431–3440.
32. Gan H, Benjamin CJ, Gibb BC (2011) Nonmonotonic assembly of a deep-cavity cavitand. J Am Chem Soc 133:4770–4773.
33. Riplinger C, Neese F (2013) An efficient and near linear scaling pair natural orbital based local coupled cluster method. J Chem Phys 138:1–18.
34. Gibb CLD, Gibb BC (2014) Binding of cyclic carboxylates to octa-acid deep-cavity cavitand. J Comput Aided Mol Des 28:319–325.
35. Sullivan MR, Sokkalingam P, Nguyen G, et al (2017) Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. J Comput Aided Mol Des 31:21–28.
36. Hanwell MD, Curtis DE, Lonie DC, et al (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminform 4:17.
37. Addicoat MA, Vankova N, Akter IF, Heine T (2014) Extension of the universal force field to metal-organic frameworks. J Chem Theory Comput 10:880–891.
38. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 97:10269–10280.
39. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) A new general purpose quantum mechanical molecular model J Am Chem Soc 107:3902–3909.
40. Besler BH, Merz KM, Kollman PA (1990) Atomic charges derived from semiempirical methods. J Comput Chem 11:431–439.
41. Frisch MJ, Trucks GW, Schlegel HB, et al Gaussian~09 {R}evision {E}.01.
42. Seminario JM (1996) Calculation of Intramolecular Force Fields from Second-Derivative Tensors. Int J Quantum Chem 30:1271–1277.
43. Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. Phys Rev Lett 91:146401.
44. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. Phys Chem Chem Phys 7:3297–305.
45. Furche F, Ahlrichs R, Hättig C, et al (2014) Turbomole. Wiley Interdiscip Rev Comput Mol Sci 4:91–100.
46. Nilsson K, Lecerof D, Sigfridsson E, Ryde U (2003) An automatic method to generate force-field parameters for hetero-compounds. Acta Crystallogr - Sect D Biol Crystallogr 59:274–289.
47. Case DA, Cerutti TE, Cheatham, III TE, et al (2017) Amber 16.
48. Wang JM, Wolf RM, Caldwell JW, et al (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174.
49. Jorgensen WL, Chandrasekhar J, Madura JD, et al (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 79:926–935.
50. Genheden S, Ryde U (2011) A comparison of different initialization protocols to obtain statistically independent molecular dynamics simulations. J Comput Chem 32:187–195.
51. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341.
52. Wu X, Brooks BR (2003) Self-guided Langevin dynamics simulation method. Chem Phys Lett 381:512–518.
53. Berendsen HJC, Postma JPM, van Gunsteren WF, et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690.
54. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N -log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092.
55. Wang K, Sokkalingam P, Gibb BC (2016) ITC and NMR analysis of the encapsulation of fatty acids within a water-soluble cavitand and its dimeric capsule. Supramol Chem 28:84–90.
56. Steinbrecher T, Mobley DL, Case DA (2007) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. J Chem Phys 127:1–1.
57. Steinbrecher T, Joung I, Case DA (2011) Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations. J Comput Chem 32:3253–3263.9
58. Gilson MK, Given JA, Bush BL, Mccammon JA (1997) The Statistical-Thermodynamic Basis for Computation of Binding Affinities : A Critical Review. 72:
59. Genheden S, Nilsson I, Ryde U (2010) Binding affinities of factor Xa inhibitors estimated by rigorous and simplified methods. J Chem Inf Model 51:947–958
60. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129:
61. Stewart JJP (2007) Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. J Mol Model 13:1173–1213.
62. Korth M (2010) Third-generation hydrogen-bonding corrections for semiempirical QM methods and force fields. J Chem Theory Comput 6:3808–3816.
63. Jurečka P, Černý J, Hobza P, Salahub DR (2007) Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with *ab initio* quantum mechanics calculations. J Comput Chem 28:555–569.
64. Antony J, Sure R, Grimme S (2015) Using dispersion-corrected density functional theory to understand supramolecular binding thermodynamics. Chem Commun 51:1764–1774.
65. Jurečka P, Černý J, Hobza P, Salahub DR (2007) Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with *ab initio* quantum mechanics calculations. J Comput Chem 28:555–569.
66. Klamt A, Schüürmann G (1993) Cosmo - a New Approach

To Dielectric Screening in Solvents With Explicit Expressions for the Screening Energy and Its Gradient. J Chem Soc - Perkin Trans 2 799–805.
67. Schäfer A, Klamt A, Sattel D, et al (2000) COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems. Phys Chem Chem Phys 2:2187–2193.
68. Klamt A (1995) Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. J Phys Chem 99:2224–2235.
69. Eckert F, Klamt A (2002) Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. AIChE J 48:369–385.
70. Eckert F, Klamt A (2010) COSMOtherm.
71. Becke AD (1988) Density-Functional Exchange-Energy Approximation With Correct Asymptotic-Behavior. Phys Rev A 38:3098–3100.
72. Perdew JP (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. Phys Rev B 33:8822–24.
73. Schäfer A, Horn H, Ahlrichs R (1992) Fully optimized contracted Gaussian basis sets for atoms Li to Kr. J Chem Phys 97:2571–2577.
74. Jensen F (2017) Introduction to Computational Chemistry, 3rd ed. John Wiley & Sons, Ltd, Chichester
75. Ryde U (1996) The coordination of the catalytic zinc in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. J Comput Aided Mol Des 10:153–164.
76. Ryde U, Olsson MHM (2001) Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. Int J Quantum Chem 81:335–34.7
77. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132:154104 (19 pages).
78. Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. J Comput Chem 32:1456–1465
79. Eichkorn K, Treutler O, Öhm H, et al (1995) Auxiliary basis-sets to approximate coulomb potentials. Chem Phys Lett 240:283–289.
80. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. Theor Chem Acc 97:119–124.
81. Sierka M, Hogekamp A, Ahlrichs R (2003) Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. J Chem Phys 118:9136–9148.
82. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129:124105.
83. Kaus JW, Pierce LT, Walker RC, Mccammon JA (2013) Improving the Efficiency of Free Energy Calculations in the Amber Molecular Dynamics Package. J Chem Theory Comput 9:4131–4139
84. Genheden S, Ryde U (2010) How to obtain statistically converged MM/GBSA results. J Comput Chem 31:837–846.
85. Bhattacharyya A (1943) On A Measure of Divergence Between Two Statistical Populations Defined by their Probability Distributions. Bull Cal Math Soc 35:99–109
86. Wu D, Kofke DA (2005) Phase-space overlap measures. I. Fail-safe bias detection in free energies calculated by molecular simulation. J Chem Phys 123:1–10.
87. Rod TH, Ryde U (2005) Quantum mechanical free energy barrier for an enzymatic reaction. Phys Rev Lett 94:138302 (4 pages).
88. Mikulskis P, Genheden S, Ryde U (2014) A large-scale test of free-energy simulation estimates of protein-Ligand binding affinities. J Chem Inf Model 54:2794–2806.
89. Sun H, Gibb CLD, Gibb BC (2008) Calorimetric Analysis of the 1:1 Complexes Formed between a Water-soluble Deep-cavity Cavitand, and Cyclic and Acyclic Carboxylic Acids. Supramol Chem 20:141–147.
90. Ponder JW, Wu C, Pande VS, et al (2010) Current Status of the AMOEBA Polarizable Force Field. J Phys Chem B 114:2549–2564.
91. Marenich A V., Cramer CJ, Truhlar DG (2009) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. J Phys Chem B 113:6378–6396.

Table 1. Raw calculated relative binding free energies (kJ/mol) for the OAH and OAM hosts, obtained with FEP at the MM and QM/MM levels for the perturbation scheme in Figure 2.

		$\Delta\Delta G_{\mathrm{bind}}^{\mathrm{MM}}$		$\Delta\Delta G_{\mathrm{bind}}^{\mathrm{QM/MM}}$	
		OAH	OAM	OAH	OAM
A5 → A6		-14.4±0.1	-16.1±0.3	-14.2±0.5	-17.7±0.5
A6 → A7		-4.3±0.2	-7.7±0.2	-6.2±0.5	-8.6±0.5
	entire ligand	0.1±0.8		-1.8±0.9	
A7 → A8		-7.1±0.3	-8.9±0.1	-7.2±0.5	-8.1±0.5
	entire ligand	-7.9±0.8		-8.0±1.0	
G0 → A7		-1.8±0.4	-9.0±0.5	-2.1±0.6	-5.8±0.7
G1 → A6		-5.8±0.3	-2.3±0.3	-11.2±0.5	-8.0±1.0
G2 → A7		16.8±0.4	7.7±0.4	6.1±0.5	1.3±1.3
G2 → MeHx		-4.2±1.5	4.3±2.6		
G3 → A6		-8.7±0.4	-10.9±0.4	-7.0±0.5	-8.9±0.6
G4 → A8		3.9±2.1	3.6±1.9	1.2±2.2	3.4±1.9
G5 → A5		2.9±0.3	-1.6±0.4	2.1±0.6	-1.2±0.5
G5 → G7		-10.2±0.2	-6.9±0.3	-6.4±0.5	-4.3±0.6
G6 → A5		9.4±0.2	9.3±0.2	8.4±0.5	9.1±0.5
A6@OAM→A6@OAH		3.6±0.4		5.2±0.6	

Table 2. Calculated absolute binding free energies (kJ/mol) for the SAMPL6 ligands in the OAH and OAM hosts obtained with FEP at the MM and QM/MM levels. The absolute affinities were obtained by using experimental data for A6, A8 or MeHx bound to OAH [34, 55]. The reference employed is specified in the columns Ref. The experimental data are given in the Exp. columns. The last nine rows show quality measures compared to the experimental results.

	OAH				OAM			
	Ref.	Exp.	MM	QM/MM	Ref.	Exp.	MM	QM/MM
A5	A6		-7.4±0.1	-7.7±0.5	A6		-9.3±0.5	-9.3±0.7
A6		-21.8±0.1			A6		-25.4±0.4	-27.0±0.6
A7	A8		-21.0±0.3	-20.9±0.5	A6		-28.8±1.0	-31.3±1.1
A8		-28.0±0.1			A6		-37.7±1.0	-39.4±1.1
G0	A8	-23.8±0.1	-19.2±0.5	-18.8±0.7	A6	-25.4±0.1	-19.8±1.1	-25.5±1.2
G1	A6	-19.5±0.1	-16.1±0.1	-10.6±0.5	A6	-25.0±0.2	-23.1±0.5	-19.0±1.1
G2	A8[a]	-35.1±0.1	-27.6±0.8	-27.0±0.6	A6	-28.5±0.1	-36.5±1.0	-32.6±1.6
G3	A6	-21.7±0.1	-13.1±0.1	-14.9±0.5	A6	-23.4±0.2	-14.5±0.6	-18.1±0.7
G4	A8	-29.7±0.1	-31.9±2.1	-29.2±2.2	A6	-32.6±0.1	-41.3±2.1	-42.8±2.2
G5	A6	-19.2±0.1	-10.3±0.3	-9.8±0.6	A6	-17.4±0.1	-7.7±0.5	-8.1±0.7
G6	A6	-20.8±0.1	-16.8±0.3	-16.1±0.5	A6	-22.6±0.1	-18.6±0.5	-18.4±0.7
G7	A6	-26.0±0.1	-20.5±0.4	-16.2±0.6	A6	-17.3±0.1	-14.6±0.6	-12.4±0.8
MeHx		-31.8±0.3			A6		-32.2±2.8	
MAD			5.6±0.3	6.7±0.3			6.2±0.3	5.5±0.4
MADtr			2.6±0.3	2.4±0.4			5.2±0.4	5.0±0.5
MSD			5.0±0.3	6.7±0.3			2.0±0.4	1.9±0.4
RMSD			6.0±0.2	7.3±0.2			6.8±0.4	6.2±0.5
Max			8.9±0.3	9.8±0.5			9.7±0.9	10.2±1.6
slope			1.1±0.1	1.1±0.1			2.0±0.1	2.1±0.1
R^2			0.77±0.05	0.81±0.04			0.85±0.02	0.93±0.02
τ			0.79±0.02	0.79±0.06			0.71±0.05	0.86±0.07
τ_{90}			0.79±0.02	0.84±0.02			0.84±0.02	1.00±0.01

[a] MeHx for MM-FEP

Table 3. Calculated relative binding free energies (kJ/mol) for the SAMPL6 ligands in the OAH and OAM hosts obtained with FEP at the MM and QM/MM levels. The relative affinities involving only the SAMPL6 ligands were obtained by using 1–3 perturbations from Table 1 and the intermediate ligands are specified in the second column. The experimental results for the SAMPL6 ligands are given in the Exp. columns.

		OAH			OAM		
OAH	Via	Exp.	MM	QM/MM	Exp.	MM	QM/MM
G0 → G2	A7	-11.3±0.2	-18.6±0.4	-8.2±0.6	-3.1±0.1	-16.7±0.5	-7.1±1.3
G1 → G3	A6	-2.2±0.1	2.9±0.2	-4.2±0.5	1.5±0.2	8.7±0.4	0.9±1.1
G4 → G2	A8, A7	-5.3±0.1	-5.8±2.1	2.2±2.2	4.1±0.1	4.8±1.9	10.2±2.2
G5 → G6	A5	-1.6±0.1	-6.5±0.4	-6.3±0.6	-5.2±0.2	-10.9±0.2	-10.3±0.5
G5 → G7		-6.8±0.1	-10.2±0.3	-6.4±0.5	0.1±0.1	-6.9±0.3	-4.3±0.6
G0 → G1	A7, A6	4.3±0.2	3.9±0.9	10.9±1.0	0.4±0.2	1.0±0.5	10.8±1.1
G5 → G3	A5, A6	-2.5±0.1	-2.8±0.4	-5.1±0.6	-6.0±0.2	-6.7±0.5	-10.0±0.7
MAD			3.1±0.2	3.9±0.4		5.1±0.2	4.9±0.4
MSD			-1.7±0.4	1.2±0.4		-2.6±0.3	-0.2±0.5
RMSD			4.1±0.2	4.5±0.6		6.7±0.2	5.6±0.5
max			7.3±0.4	7.5±1.6		13.6±0.5	10.5±1.1
slope			1.4±0.1	0.9±0.1		2.0±0.1	2.0±0.1
R^2			0.87±0.02	0.56±0.08		0.61±0.04	0.73±0.04
τ_r			0.71±0.01	0.71±0.11		0.71±0.13	0.71±0.16
$\tau_{r,90}$			0.71±0.01	1.00±0.00		1.00±0.09	1.00±0.06

Table 4. Calculated MM → QM/MM free energies (kJ/mol) for ligands G0–G7 and A5–A8 ($\Delta\Delta G_{\mathrm{bind,L}}^{\mathrm{MM\rightarrow QM/MM}} = \Delta\Delta G_{\mathrm{L,bound}}^{\mathrm{MM\rightarrow QM/MM}} - \Delta\Delta G_{\mathrm{L,free}}^{\mathrm{MM\rightarrow QM/MM}}$). The last two columns show the number of Λ values used in the calculations.

	$\Delta\Delta G_{\mathrm{bind,L}}^{\mathrm{MM\rightarrow QM/MM}}$		#Λ	
	OAH	OAM	OAH	OAM
G0	-4.7±0.3	-8.8±0.2	4	4
G1	2.3±0.3	1.0±0.9	5	5
G2	5.7±0.4	0.8±1.2	5	6
G3	-4.8±0.3	-6.8±0.2	4	4
G4	-2.4±0.3	-4.5±0.3	5	6
G5	-2.6±0.3	-3.5±0.2	4	4
G6	-2.3±0.3	-2.9±0.2	4	4
G7	1.2±0.3	-1.0±0.2	5	4
A5	-3.3±0.3	-3.1±0.3	4	4
A6	-3.1±0.3	-4.7±0.3	4	4
A7	-5.0±0.3	-5.6±0.3	4	4
A8	-5.1±0.3	-4.8±0.3	4	4

Table 5. $\Delta\Delta G_{\mathrm{bind}}$ (kJ/mol) calculated for the SAMPL4 ligands with SQM approach and three different ways to combine the ten energies from different snapshots, plain average, the minimum energy (Min) or the Boltzmann average (Boltz). The second column show the experimental values [34].

	Exp.	Average	Min	Boltz
Bz	-15.6±0.2	-0.3±3.2	-9.2	-7.6
MeBz	-24.5±0.5	-10.4±1.5	-17.6	-15.6
EtBz	-26.2±0.1	-14.6±2.0	-23.6	-22.8
pClBz	-28.1±0.1	-17.8±2.5	-30.1	-28.9
Hx	-23.5±0.3	-11.7±2.8	-26.0	-25.3
MeHx	-31.8±0.3	-16.8±1.4	-20.5	-19.2
Pen	-15.6±0.2	-8.2±2.3	-22.2	-21.6
Hep	-27.7±0.1	-11.6±3.1	-21.0	-20.4
A8	-28.0±0.1	-26.1±2.9	-34.9	-33.5
A10	-31.5±0.1	-23.3±1.8	-30.1	-28.2
MAD		11.2±0.7	5.3	5.7
MADtr		3.4±0.6	5.1	5.2
MSD		11.2±0.8	1.7	2.9
RMSD		11.9±0.8	6.1	6.7
max		16.1±2.0	11.3	12.6
slope		1.1±0.1	0.7	0.7
R^2		0.65±0.10	0.34	0.31
τ		0.64±0.09	0.33	0.29
τ_{90}		0.92±0.03	0.38	0.33

Table 6. $\Delta\Delta G_{bind}$ (kJ/mol) calculated for the SAMPL5 ligands with SQM approach and three different ways to combine the ten energies from different snapshots, plain average, the minimum energy (Min) or the Boltzmann average (Boltz). The second column show the experimental values.

Ligand	Host	Exp	Average	Min	Boltz
S5-G1	OAH	-21.09±0.04	-2.9±2.4	-13.5	-11.8
S5-G2		-17.78±0.04	-8.6±1.9	-19.1	-17.6
S5-G4		-39.20±0.01	-22.4±3.0	-34.8	-33.4
S5-G6		-22.31±0.02	-17.2±2.0	-26.7	-25.8
S5-G1	OAM	-21.92±0.21	-10.3±1.6	-19.5	-18.5
S5-G2		-21.09±0.13	-18.6±5.9	-36.2	-35.1
S5-G4		-9.96±0.08	0.0±2.4	-13.1	-12.1
S5-G6		-18.91±0.08	-19.8±2.0	-28.7	-27.0
MAD			9.3±0.9	6.0	5.8
MADtr			5.1±0.9	5.7	5.8
MSD			9.1±1.1	-2.4	-1.1
RMSD			11.0±0.9	7.4	7.2
max			18.2±2.1	15.1	14.0
slope			0.7±0.1	0.7	0.7
R^2			0.45±0.12	0.39	0.38
τ			0.48±0.10	0.48	0.41
τ_{90}			0.65±0.03	0.48	0.41

Table 7. $\Delta\Delta G_{bind}$ (kJ/mol) calculated for the SAMPL6 ligands with SQM and DFT approaches and three different ways to combine the ten energies from different snapshots, plain average, the minimum energy (Min) or the Boltzmann average (Boltz). Values in brackets for the quality measures of DFT are calculated after subtraction of 20.9 kJ/mol to all values (used for the submitted data).

	OAH				OAM			
	SQM			DFT	SQM			DFT
	Av	Min	Boltz		Av	Min	Boltz	
G0	-10.8±2.4	-25.0	-24.4	5.0	-23.5±1.3	-37.0	-35.3	18.6
G1	-10.1±3.2	-26.9	-25.9	5.8	-18.2±1.7	-30.5	-28.6	3.2
G2	-17.1±3.9	-22.4	-21.5	-20.5	-14.7±5.3	-44.5	-43.2	5.2
G3	-6.1±2.3	-17.9	-16.6	3.5	-15.2±1.0	-21.1	-19.3	17.8
G4	-29.9±4.4	-45.6	-44.1	-27.0	-41.0±1.3	-50.8	-48.0	-22.3
G5	-7.3±2.2	-17.7	-16.4	-1.0	-16.9±2.8	-35.2	-33.6	18.5
G6	-7.5±3.1	-17.8	-16.7	13.2	-19.9±1.2	-29.6	-27.8	16.9
G7	-12.5±2.0	-22.2	-20.4	2.0	-8.7±6.1	-42.3	-41.4	12.6
MAD	11.9±1.0	6.1	6.6	22 (8.0)	6.4±1.0	12.9	11.7	33 (14.6)
MADtr	3.6±0.9	6.1	6.3	7.7	5.1±0.9	6.9	7.0	7.5
MSD	11.8±1.1	0.0	-1.2	22 (1.2)	4.3±1.2	-12.4	-10.7	33 (11.9)
RMSD	12.9±1.0	8.0	8.0	24 (9.3)	7.6±1.4	14.8	13.5	34 (15.5)
max	18.0±2.7	15.8	14.4	34 (18.2)	13.8±4.2	25.1	24.2	44 (23.1)
slope	1.0±0.2	0.7	0.3	2.0	1.4±0.3	0.8	0.7	2.0
R^2	0.51±0.16	0.18	0.18	0.66	0.55±0.15	0.17	0.14	0.57
τ	0.64±0.15	0.43	0.36	0.43	0.43±0.18	0.36	0.36	0.29
τ_{90}	0.83±0.07	0.43	0.36	0.43	0.86±0.06	0.41	0.41	0.33

Paper V

Martin A. Olsson, Alfonso. T. García-Sosa, Ulf Ryde, *Journal of Computer-Aided Molecular Design,* **2018**, 32, 221–224

J Comput Aided Mol Des
DOI 10.1007/s10822-017-0056-z

Binding affinities of the farnesoid X receptor in the D3R Grand Challenge 2 estimated by free-energy perturbation and docking

Martin A. Olsson[1] · Alfonso T. García-Sosa[2] · Ulf Ryde[1]

Received: 2 June 2017 / Accepted: 29 August 2017

Abstract We have studied the binding of 102 ligands to the farnesoid X receptor within the D3R Grand Challenge 2016 blind-prediction competition. First, we employed docking with five different docking software and scoring functions. The selected docked poses gave an average root-mean-squared deviation of 4.2 Å. Consensus scoring gave decent results with a Kendall's τ of 0.26 ± 0.06 and a Spearman's ρ of 0.41 ± 0.08. For a subset of 33 ligands, we calculated relative binding free energies with free-energy perturbation. Five transformations between the ligands involved a change of the net charge and we implemented and benchmarked a semi-analytic correction (Rocklin et al., J Chem Phys 139:184103, 2013) for artifacts caused by the periodic boundary conditions and Ewald summation. The results gave a mean absolute deviation of 7.5 kJ/mol compared to the experimental estimates and a correlation coefficient of $R^2 = 0.1$. These results were among the four best in this competition out of 22 submissions. The charge corrections were significant (7–8 kJ/mol) and always improved the results. By employing 23 intermediate states in the free-energy perturbation, there was a proper overlap between all states and the precision was 0.1–0.7 kJ/mol. However, thermodynamic cycles indicate that the sampling was insufficient in some of the perturbations.

Keywords Ligand binding · Docking · Quantum-polarised ligand docking · Free-energy perturbation · Bennett acceptance ratio · Periodic boundary conditions · Charge transformations · Drug design data resource · D3R Grand Challenge 2016

Introduction

The increase in computer power and advances in protein crystallography and drug discovery during the latest decades have nourished the dream that drugs one day may be developed by computational methods [1]. One of the most important properties of a drug candidate is its binding affinity to the receptor molecule and many computational approaches are available to calculate binding affinities [2]. One of the best is alchemical free-energy perturbation (FEP) [2–4], calculating the energies by exponential averaging, thermodynamic integration, Bennett acceptance ration (BAR), multistate BAR (MBAR) or similar methods [5–8]. Being based on strict statistical-mechanics grounds, the primary limitations of FEP are the force-field employed and the sampling of the conformational space. Several recent large-scale retrospective benchmark studies have indicated that relative binding free energies of drug-like molecules to protein targets can be calculated by FEP with a mean absolute deviation (MAD) from experimental affinities of 4–6 kJ/mol [9–12]. A similar accuracy has also been reported for prospective calculations of binding affinities in host–guest systems [13, 14]. However, for protein systems, prospective predictions have typically been quite poor with MADs of 4–16 kJ/mol [15, 16], probably owing to uncertainties and variations in the binding mode.

Large-scale studies of FEP-calculated relative binding affinities have in general been restricted to charge-preserving

Electronic supplementary material The online version of this article (doi:10.1007/s10822-017-0056-z) contains supplementary material, which is available to authorized users.

✉ Ulf Ryde
Ulf.Ryde@teokem.lu.se

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

[2] Institute of Chemistry, University of Tartu, Ravila 14a, 50411 Tartu, Estonia

Published online: 06 September 2017

transformations [9–16]. The reason for this is that perturbations of the net charge suffer from known artifacts in the treatment of electrostatics during molecular simulations with periodic boundary conditions and Ewald summation [17, 18], with the effect that the results depend on the size of the simulated periodic box and the software employed [19, 20]. In addition, a change in the net charge of the ligand gives rise to large and long-ranged electrostatic effects of the surrounding protein that may be hard to estimate accurately [20]. Many schemes have been suggested to correct FEP calculations for artifacts caused by the periodicity and the Ewald summation [17, 21–23]. However, they have been primarily directed towards solvation free energies of simple ions, often providing complicated and software-specific corrections. Recently, Rocklin et al. [24] and Reif and Oostenbrink [25] independently suggested general procedures to correct FEP predictions of relative binding free energies. Considering that many drug-design projects involve molecules with a varying net charge, it is important to test and calibrate methods that can handle such ligand series.

In this paper, we study the binding of 102 inhibitors to the farnesoid X receptor (FXR) [26] from the blind-prediction drug-design data resource (D3R) Grand Challenge 2016 (GC2) [27]. FXR, also known as the bile-acid receptor or nuclear receptor 1H4, has recently appeared as an interesting drug-discovery target, providing an alternative to surgical treatment of obesity [28]. The binding site of FXR is located between two flexible α-helices, such that the ligands are typically pinched between residues His-298 and Met-294 [26, 29]. This flexible binding makes FXR a challenging target for computational approaches. Moreover, the inhibitors have a varying net charge, 0 or −1. We have studied these inhibitors with two set of methods. First, we have tried to estimate the binding mode and binding affinities for all 102 ligands with five different docking and scoring methods. Second, for a subset of 33 ligands, we have tried to provide more accurate relative binding affinities by employing FEP methods. To this end, we have implemented the approach of Rocklin et al. [24] in combination with the AMBER software [30] to provide corrections for ligand transformations that involve a change in the net charge of the ligand. Thereby, we obtain a prospective benchmark test of this approach in a real drug-design problem. Furthermore, we thoroughly asses the results in terms of overlap criteria and thermodynamic cycles.

Methods

Protein setup

Three crystal structures were employed in our calculations. The starting structure for the docking calculations was the 3OMK structure [29], because it had the highest resolution among the available crystal structures, 1.9 Å. Moreover, it contained a benzimidazole ligand that resembled some of the challenge ligands and the binding site was large enough to accommodate all the ligands in the set. For the FEP simulations, we employed crystal structures of FXR complexed with ligands **12** and **17**, provided by the GC2 organisers in the second stage of the challenge. All structures were prepared and hydrogen atoms were added using the protein preparation wizard in the Schrödinger Maestro software [31], assuming a pH of 7.4, employed in the binding assay [27]. We also analysed possible hydrogen-bond interactions, the solvent exposure and the local surroundings of the histidine residues by local software [32] and visual inspection. Based on this analysis, we concluded that His-317, 426, 449 and 450 are protonated on the ND1 atom, whereas the other six His residues (two of which are in the ligand-binding site, His-298 and 451) are protonated on the NE2 atom. His-449 and 450 were flipped (i.e. the C and N atoms in the imidazole ring were interchanged). All water molecules were kept in the calculations.

Docking and scoring

Before the docking, the 3OMK structure without the ligand was solvated in an octahedral box of TIP4P-Ew water molecules [33] extending at least 10 Å from the solute and was equilibrated by molecular dynamics (MD) for 10 ns. The distance between residues His-298 and Met-294 was monitored (Fig. S1) and the snapshot with the largest distance was selected for the docking (giving the most open binding site), because initial docking calculations suggested that some of the ligands were too large for the binding site.

Five docking approaches were used: Schrödinger quantum-polarised ligand docking (QPLD [34], v. 2016), Glide SP (single precision), Glide XP (extended precision) [35], AutoDock4 [36] and AutoDock Vina (Vina, version 1.12) [37], which employ different algorithms and/or scoring functions. Ligand conformational libraries were generated using LigPrep [38]. Preparation for docking with Vina was done using MGLTools [39] with ligand files from LigPrep. We employed a larger than default exhaustiveness of global search (exhaustiveness = 12, rather than 8).

For the final scoring, two different methods were used. In the first, a consensus score (CS) was employed involving the average of the five scores from QPLD, Glide SP, Glide XP, AutoDock 4 and Vina. This was done in order to hedge predictions from unreasonably high or low values. In the second, the same scoring functions were used, but the average of the ranks was used, instead of the scores (CR, consensus rank).

Free-energy simulations

Two sets of ligands for FEP calculations (FEP sets 1 and 2) were included in the GC2, involving 33 ligands in total. We use the numerical names of the ligands, suggested by the GC2 organisers, which are shown in Fig. 1. The FEP simulations were started from crystal structures of FXR with ligands **12** and **17** for sets 1 and 2, respectively. The other ligands were built inside the active site, based on these structures using Avogadro software [40] and the geometry was optimised with the UFF force field [41].

The ligands were manually mapped for the FEP simulations, minimising the difference between the ligands and the number of perturbations changing the net charge of the ligand. The transformations are also shown in Fig. 1. In order to assess the convergence of the binding energies, cycles were introduced when possible without introducing larger perturbations than for the other transformations (five for FEP set 1 and four for set 2). In one case, this involved the addition of an extra ligand, **M1**, also shown in Fig. 1. Five of the perturbations involved a change in the net charge of the ligand and therefore required corrections to the calculated binding free energies, viz. **101→91, 88→79, 75→88, 41→12, 88→73** and **M1→84**.

All FEP simulations were performed with the AMBER 14 and 16 software [30] with the ff14SB force-field [42] for FXR and the GAFF force field [43] for the ligands. Charges for the ligands were derived by first geometry optimising the ligands at the AM1 [44] level, followed by a calculation of the electrostatic potential at the HF/6-31G* [45] level of theory at points sampled according to the Merz–Kollman scheme [46]. These calculations were performed with the Gaussian09 [47] software. Finally, restrained electrostatic-potential charges [48] were fitted to the electrostatic potential using the antechamber program in the AMBER software [30]. The Seminario approach [49] implemented in the Hess2FF program [50] was used to obtain missing torsion parameters of the ligands, based on frequency calculations performed at the BLYP/def2-SVP level of theory [51–53]. Added parameters are listed in Table S1 in the SI.

For the FEP simulations, FXR and the ligands were solvated in a truncated octahedral box of TIP3P water molecules [54], extending at least 9 Å from the solute using the leap program in the AMBER suite, so that ~8000 water molecules were surrounding the solute (for perturbations modifying the net charge of the ligand, cubic boxes were used instead, see below). TIP3P water molecules were used for the binding affinities, because they have been shown to give the best energies [55], whereas TIP4P-Ew gave better dynamical properties [56].

The FEP simulations were run with the pmemd module of AMBER, using the dual topology scheme with both ligands in the topology file [57]. Each ligand transformation was divided into steps 25 steps, employing a linear transformation of the force-field potentials with the coupling parameter $\lambda = 0.0, 0.025, 0.050, 0.075, 0.10, 0.15, 0.20, \ldots, 0.80, 0.85, 0.90, 0.925, 0.95, 0.975$ and 1.0. Electrostatic and van der Waals interactions were perturbed concomitantly, using soft-core potentials for both types of interactions [58, 59]. Soft-core potentials were used not only for atoms differing between the two ligands, but also for all atoms in the ligand ring systems neighbouring the perturbed group to allow for larger differences in the dynamics of the perturbed groups (atoms without soft-core potentials have identical coordinates in the perturbations).

For each λ value, 100 steps of minimisation were performed with the heavy atoms of the protein and ligand restrained towards the starting structure with a force constant of 418 kJ/mol/Å^2. This was followed by 20 ps constant-volume equilibration with the same restraints and 2 ns constant-pressure equilibration without any restraints. Finally, a 2 ns production simulation was run for each of the 25 λ values, during which structures and energies were sampled every 2 ps.

In all the MD and FEP simulations, bonds involving hydrogen atoms were constrained with the SHAKE algorithm [60], allowing for a time-step of 2 fs. The temperature was kept constant at 300 K using Langevin dynamics [61] with a collision frequency of 2 ps^{-1}, and the pressure was kept constant at 1 atm using a weak-coupling isotropic algorithm [62] with a relaxation time of 1 ps. Long-range electrostatics were handled by particle-mesh Ewald summation [63] with a fourth-order B spline interpolation and a tolerance of 10^{-5}. The cut-off for Lennard-Jones interactions was set to 8 Å. No counter-ions were used in the calculations.

Relative binding free energies between two ligands, L_0 and L_1 ($\Delta\Delta G_{\text{bind}}$), were estimated using a thermodynamic cycle that relates $\Delta\Delta G_{\text{bind}}$ to the free energy of alchemically transforming L_0 into L_1 when they were either bound to the protein, $\Delta\Delta G_{\text{bound}}$, or were free in solution, $\Delta\Delta G_{\text{free}}$ [64]:

$$\Delta\Delta G_{\text{bind}} = \Delta G_{\text{bind}}(L_1) - \Delta G_{\text{bind}}(L_0) = \Delta\Delta G_{\text{bound}} - \Delta\Delta G_{\text{free}} \quad (1)$$

$\Delta\Delta G_{\text{bound}}$ and $\Delta\Delta G_{\text{free}}$ were estimated by the multi-state Bennett acceptance-ratio (MBAR) method [8], using the pymbar software [8], including only statistically non-correlated energies in the calculations. For comparison, BAR energies were also employed, calculated with the same software.

Charge-transformation corrections

In this study, raw $\Delta\Delta G_{\text{bind}}$ estimates for ligand transformations that modified the net charge of the ligand were corrected for errors caused by the use of periodic boundary-conditions and Ewald summations in the FEP simulations,

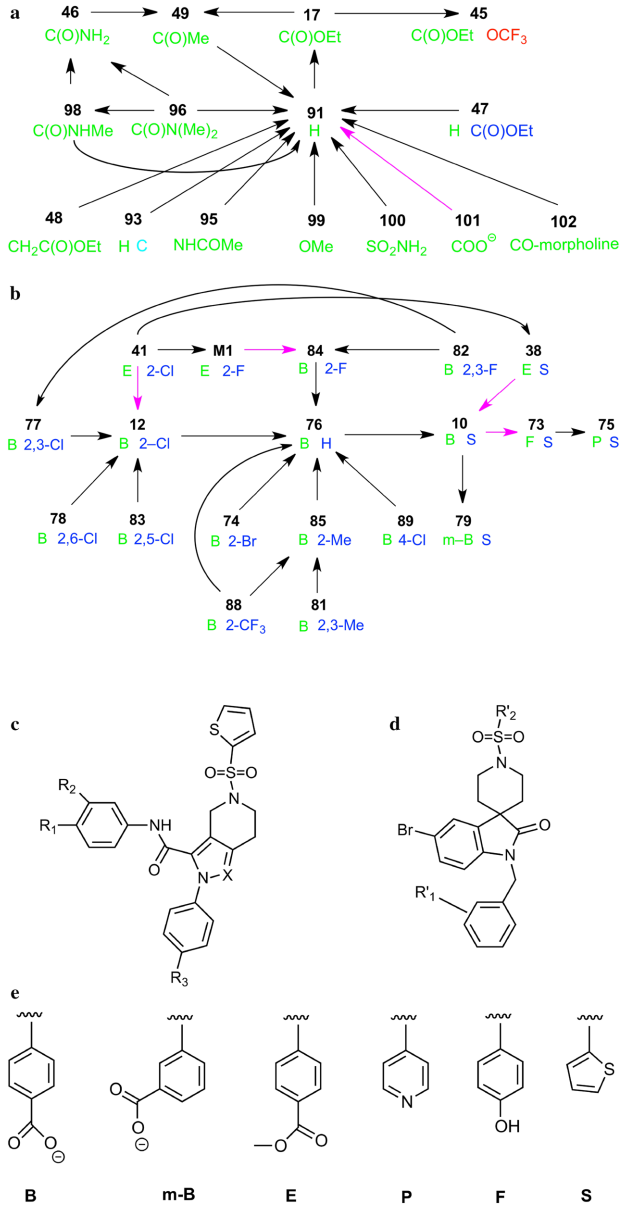


Fig. 1 Ligands transformations studied for FEP sets 1 (**a**) and 2 (**b**) (*arrows* in *magenta* indicate perturbations modifying the net charge of the ligand). **c** and **d** show the general scaffold of the ligands in sets 1 and 2, respectively. In **a**, substituents R_1, R_2 and R_3, as well as the varying atom X are shown in *green, blue, red and cyan*, respectively. Only R_1 is shown for all ligands, whereas the other three are shown only if they differ from –H, –H and N, respectively. In **b**, the R'_1 and R'_2 substituents are shown in *green* (*left*) and *blue* (*right*), respectively. The former is shown in *one-letter codes*, explained in **e**, whereas the latter is shown either with S, indicating a thiophene group, also shown in **e**, or a substituted benzene ring, in which case only the substituents are shown, with the numbering starting from the position connected to the remainder of the molecule

giving corrected binding free-energies ($\Delta\Delta G_{\text{bind, corr}}$) that are independent of simulated box size. We have employed the semi-analytic correction suggested by Rocklin et al. [24]. It requires that the FEP calculations are run in a cubic periodic box with a constant volume. The free energies can then be corrected by calculating the residual integrated potential

(RIP) for three non-periodic systems by numerically solving the Poisson–Boltzmann equation. All three calculations involve the protein–ligand complex with all water molecules removed. In the first calculation, the protein atoms have full charges (taken from the MM force field), whereas the ligand charges were zeroed. The other two calculations have full ligand charges but zeroed protein charges. They differ in the value of the dielectric constant of the solvent: in the first calculation (as well as in the calculation with zeroed ligand charges), the solvent dielectric constant was that of the bulk solvent ($\varepsilon_s = 97$ for TIP3P water [24, 54]). In the second calculation, the solvent dielectric constant was the same as the internal dielectric constant, which was unity in all calculations. The resulting RIPs from these three calculations will be denoted I_P, I_L and $I_{L,hom}$ below.

Based on these RIPs, five corrections to $\Delta\Delta G_{bind}$ were calculated, as was detailed by Rocklin et al. [24]: a correction for periodicity-induced net-charge interactions (ΔG_{NET}), a correction for periodicity-induced net-charge undersolvation (ΔG_{USV}), a correction for RIP effects (ΔG_{RIP}), an empirical correction to reproduce the exact analytical result in the special case of a single point charge at the centre of a spherical cavity (ΔG_{EMP}) and a correction for discrete solvent effects (ΔG_{DSC}). These five terms were calculated in the following way [24]:

$$\Delta G_{NET} + \Delta G_{USV} = -\frac{\xi_{LS}}{8\pi\epsilon_0}\frac{\left(\left(Q_P + Q_L\right)^2 - Q_P^2\right)}{\epsilon_s L} \tag{2}$$

$$\Delta G_{RIP} = \frac{\left(I_P + I_L\right)\left(Q_P + Q_L\right) - I_P Q_P}{L^3} \tag{3}$$

$$\Delta G_{EMP} = -\frac{1}{8\pi\epsilon_0}\frac{16\pi^2}{45}\left(1 - \frac{1}{\epsilon_s}\right)\left[\left(Q_P + Q_L\right)^2 - Q_P^2\right]\frac{R_L^5}{L^6} \tag{4}$$

$$\Delta G_{DSC} = -\frac{\gamma_s Q_L N_s}{6\epsilon_0 L^3} \tag{5}$$

In these equations, Q_L and Q_P are the net charge of the ligand and the protein, respectively (−1 and −10 in the present calculations), L is the side length of the cubic periodic box (~7.9 nm), N_s is the number of solvent molecules in the periodic box (~14000), ε_0 is the permittivity of vacuum, ξ_{LS} is the cubic lattice-sum (Wiegner) integration constant (–2.837), ε_s is the static relative dielectric permittivity of the solvent ($\varepsilon_s = 97$ for TIP3P water [24, 54]), γ_s is the quadrupole-moment trace of the solvent model relative to its single van der Waals interaction site, which for TIP3P is 0.00764 e nm^2 (note that ref. 24 gives a 10 times too large value) and the effective solvation radius is calculated from

$$R_L = \sqrt{\frac{I_L - I_{L,hom}}{\frac{1}{8\pi\epsilon_0}\frac{4\pi}{3}\left(1 - \frac{1}{\epsilon_s}\right)|Q_L|}} \tag{6}$$

Two sets of calculations were needed, one for the protein–ligand simulation and one for the free ligand in water solution. In the latter case, only I_L and $I_{L,hom}$ can be calculated, whereas $I_P = Q_P = 0$ in Eqs. 2–4. Corrections are needed only for the charged ligand (the terms vanish for $Q_L = 0$). The final corrected binding energy was then calculated as the sum of the original binding free energy (obtained from the simulations with periodic boundary conditions and Ewald summation) and these five correction terms (taken as the difference between the corrections obtained for the protein–ligand complex and for the free ligand):

$$\Delta\Delta G_{bind,corr} = \Delta\Delta G_{bind} + \Delta\Delta G_{NET} + \Delta\Delta G_{USV} + \Delta\Delta G_{RIP} + \Delta\Delta G_{EMP} + \Delta\Delta G_{DSC} \tag{7}$$

The Poisson–Boltzmann calculations were run by the APBS software [65], using PARSE [66] radii for all atoms. A cubic grid of 257^3 points were employed with a side length of ~80 Å for the protein–ligand complex and ~39 Å for the ligand. To ensure that the estimates are stable, the Poisson–Boltzmann calculations were performed for eight snapshots from the simulations, also allowing for an estimate of the uncertainty of the calculations. The RIPs were calculated from the APBS output by Python scripts provided by the authors of ref. [24]. We have designed a semi-automatic procedure to perform all the needed calculations, based on the AMBER FEP simulation files. The procedure and the needed files can be found in http://signe.teokem.lu.se/~ulf/Methods/ChargedFEPCorrections.html.

Uncertainties and convergence measures

All reported uncertainties are standard errors of the mean (standard deviations divided by the square root of the number of samples). The uncertainty of the MBAR free energies calculated at each λ value was estimated by bootstrapping using the pymbar software [8] and the total uncertainty was obtained by error propagation (the total variance was the sum of the individual variances).

The performance of the free-energy estimates was quantified by the mean absolute deviation (MAD), the correlation coefficient (R^2), Kendall's rank correlation coefficient (τ) and Spearman's rank correlation coefficient (ρ) compared to the experimental data from GC2 [27]. For the FEP calculations, τ was calculated only for the transformations that were explicitly studied, not for all combinations that can be formed from these transformations (τ_r). Moreover, it was also evaluated considering only differences (both experimental and calculated) that are statistically

significant at the 95% level ($\tau_{r,95}$) [67]. It should be noted that R^2 depends on the direction of the FEP perturbation (i.e. whether **12→41** or **41→12** was considered, which is arbitrary). This was solved by considering both directions when R^2 was calculated. The standard deviation of the quality measures was obtained by a simple simulation approach [68]: for each transformation, a $\Delta\Delta G_{bind}$ result was sampled as a random number from a Gaussian distribution with the mean and standard error obtained from the MBAR calculations. The quality measures were then calculated and the procedure was repeated 1000 times. The standard error of these estimates is reported as the uncertainty. Since no uncertainty in the experimental affinities was reported [27], we assumed a typical uncertainty of 2.4 (=1.7 $\sqrt{2}$) kJ/mol [69] for these values when estimating the precision of the quality measures.

To assess the convergence of the various FEP calculations, seven overlap measures were employed [10]: the Bhattacharyya coefficient for the energy distribution overlap (Ω), the Wu & Kofke overlap measures of the energy probability distributions (K_{AB}), as well as their bias metrics (Π), the weight of the maximum term in the exponential average (w_{max}), the difference between the forward and backward exponential average estimate ($\Delta\Delta G_{EA}$), the difference between the MBAR and BAR estimates ($\Delta\Delta G_{BAR}$) and the standard deviation of the energies (σ) [10, 70–72]. Moreover, the reliability of the free-energy estimates was assessed by adding cycles among the FEP transformation, as is shown in Fig. 1. The cycle-closure hystereses give an estimate of the errors from incomplete sampling of the phase space.

Results and discussion

As a part of the D3R Grand Challenge 2016, we have performed a prospective study of the binding of 102 inhibitors to FXR. We employed two sets of calculations: docking and scoring with five different software or scoring functions, and FEP calculations for the two FEP subsets, involving semi-analytic corrections [24] for the change in the net charge of some ligand pairs. The results of these calculations are described in separate sections.

Docking results

102 rather diverse ligands, most of them belonging to four chemical motifs, benzimidazole, isoxazole, spiro and sulfonamides, were docked to FXR. As mentioned in the "Methods" section, we employed five different docking approaches: QPLD, Glide SP, Glide XP, AutoDock 4 and AutoDock Vina. The submitted poses were those with the lowest energy from QPLD, because we expected that this method would give the most accurate results [34, 73] (the ligand charges are polarised by the surrounding protein). In six cases, QPLD did not provide any acceptable pose (a pose that fitted into the binding site). In those cases, we used instead either the Glide XP pose if acceptable (**16**) or Vina poses (**65, 79, 80, 97** and **101**).

After the results were submitted, crystal structures of FXR with 35 of the ligands were revealed. Our docked ligand binding poses were in line with those of the other submissions. It should be noted that we submitted only a single pose for each complex, whereas most other submissions involved more than one predicted binding pose. In several cases, reasonable poses were obtained, as can be seen in Table 1 (last column). Predictions with a root-mean-squared deviation (RMSD) from the crystal structures of 2 Å or less were obtained for 16 of the ligands (46%; **7, 13, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32** and **36**). The average RMSD for all structure predictions was 4.2 Å, which puts our results at position 22 among the 51 complete submissions for pose predictions. The best result (RMSD = 1.1 Å) was obtained for ligand **28**, which is shown in Fig. 2a. The largest RMSD was 9.6 Å for **34**, shown in Fig. 2b, for which the docking failed to reproduce the extended conformation of the ligand in the crystal structure.

Our two scoring functions, CS and CR, (submitted before the crystal structures were revealed) gave nearly identical results compared to the experimental affinities [27]: Kendall's τ was 0.26 ± 0.06 for both, whereas the Spearman's ρ was 0.40 ± 0.09 and 0.41 ± 0.08, respectively, as calculated by the GC2 organisers. These results were in the middle among the submissions, at positions 34 and 35, respectively, out of 59 submissions.

For simplicity, we submitted only one docking pose, the one with the lowest score. Different methods can be devised to use more than one pose, e.g. by combining the scores from several poses or by providing several poses with varying scores. Given that our procedure also included several docking programs and their different algorithms and scoring functions, we decided to use the consensus ranks and scores. Other protein crystal structures may have been used, but we found the structure chosen suitable for the task. The flexibility and dynamics of the binding site and ligands may also have been explored, but the given time was not enough for a deeper study.

FEP results

We have estimated the relative binding affinities of 33 ligands of FXR by FEP calculations with the AMBER software. The ligands were divided by the organiser into two sets: FEP set 1, involving 18 sulfonamide ligands, and FEP set 2, with 15 spiro ligands. We set up two networks involving 19 and 20 transformations for the two sets, respectively,

Table 1 Results of the docking calculations with five software and scoring functions (Glide XP, Glide SP, AutoDock 4, AutoDock Vina and QPLD)

Ligand	XP	SP	AD4	Vina	QPLD	CS	CS rank	CR rank	RMSD
1	−35.6	−32.6	−38.2	−36.8	−38.7	−36.4	53	70	7.7
2	−24.0	−33.3	−36.9	−37.7	−37.0	−33.8	71	79	7.4
3	−35.8	−33.7	−42.6	−45.2	−41.2	−39.7	46	50	6.4
4	−37.1	−26.8	−45.0	−39.7	−40.9	−37.9	50	54	7.0
5	−37.5	−31.7	−33.6	−35.1	−38.6	−35.3	59	73	7.3
6	−49.2	−43.7	−48.2	−50.2	−49.0	−48.1	14	13	6.7
7	−55.9	−49.0	−53.5	−55.6	−58.6	−54.5	1	1	1.2
8	−47.5	−42.6	−42.8	−51.5	−39.0	−44.7	29	35	6.5
9	−36.1	−41.2	−49.1	−54.8	−40.0	−44.2	30	33	6.9
10	−33.8	−33.6	−25.0	−26.8	−38.6	−31.6	80	92	3.3
11	−26.3	−31.6	−26.0	−22.2	−37.6	−28.7	89	99	9.4
12	−10.0	−35.7	−26.3	−24.7	−41.0	−27.5	92	83	3.4
13	−41.8	−45.4	−52.5	−58.6	−41.7	−48.0	15	11	1.3
14	−36.9	−42.8	−47.2	−53.1	−38.8	−43.8	35	40	6.5
15	−43.3	−26.8	−42.3	−38.9	−40.6	−38.4	48	52	6.3
16	−29.9	−26.5	−32.3	−29.3	−16.7	−26.9	95	95	6.0
17	−39.4	−37.2	−32.4	−24.7	−41.6	−35.1	63	59	6.2
18	−51.0	−36.9	−39.2	−44.8	−41.5	−42.7	41	38	9.3
19	−51.0	−42.9	−48.1	−51.5	−54.6	−49.6	10	12	1.3
20	−56.1	−44.2	−49.0	−50.2	−59.1	−51.7	5	6	1.2
21	−54.7	−44.0	−46.9	−55.2	−56.6	−51.5	6	5	1.2
22	−52.2	−41.6	−45.9	−50.2	−54.2	−48.8	11	14	1.7
23	−46.7	−38.5	−47.1	−42.3	−41.7	−43.3	39	30	4.5
24	−52.8	−46.7	−46.4	−57.3	−56.6	−52.0	4	3	1.6
25	−56.2	−41.3	−49.9	−56.5	−59.2	−52.6	2	2	1.2
26	−54.1	−39.3	−50.0	−53.6	−56.3	−50.7	9	9	1.4
27	−54.9	−52.0	−39.7	−51.5	−40.3	−47.7	18	18	1.3
28	−55.9	−50.2	−42.8	−54.4	−40.3	−48.7	12	10	1.1
29	−55.4	−48.6	−42.0	−53.6	−40.4	−48.0	16	15	1.4
30	−38.7	−45.4	−34.5	−44.8	−40.0	−40.7	45	49	2.0
31	−51.5	−48.3	−45.7	−49.4	−41.6	−47.3	19	15	1.8
32	−56.7	−43.0	−38.8	−40.2	−56.1	−47.0	21	20	1.4
33	−35.5	−16.7	−39.4	−23.8	−40.7	−31.2	82	71	
34	−16.7	−17.4	−22.6	−25.9	−16.7	−19.9	101	101	9.6
35	−40.4	−47.1	−34.4	−43.5	−41.1	−41.3	44	43	2.5
36	−55.8	−53.8	−20.1	−42.7	−40.0	−42.5	42	41	1.8
37	−41.1	−41.0	−41.7	−43.5	−41.0	−41.7	43	42	
38	−37.2	−23.8	−34.7	−26.8	−41.8	−32.8	75	62	
39	−45.5	−43.8	−47.3	−54.0	−39.2	−46.0	24	23	
40	−49.0	−40.8	−43.8	−53.1	−38.1	−45.0	28	37	
41	−40.3	−35.4	−29.9	−23.4	−39.0	−33.6	73	77	
42	−47.9	−40.3	−37.6	−53.1	−41.8	−44.1	31	26	
43	−36.7	−23.2	−41.2	−37.2	−39.8	−35.6	58	60	
44	−52.1	−41.6	−42.4	−41.8	−41.0	−43.8	33	29	
45	−16.7	−16.7	−33.1	−15.9	−36.4	−23.8	98	100	
46	−38.7	−37.7	−38.7	−23.8	−41.0	−36.0	55	56	
47	−39.6	−36.6	−40.1	−31.4	−41.8	−37.9	51	48	
48	−43.8	−36.7	−32.4	−25.1	−41.0	−35.8	57	57	
49	−40.0	−35.8	−38.5	−21.8	−38.0	−34.8	64	72	
50	−52.0	−41.5	−42.0	−48.1	−51.8	−47.1	20	21	
51	−55.4	−39.9	−49.3	−52.3	−59.6	−51.3	7	8	
52	−53.3	−49.2	−48.7	−51.5	−58.3	−52.2	3	4	
53	−49.7	−50.3	−37.4	−52.3	−40.0	−45.9	25	27	

Table 1 (continued)

Ligand	XP	SP	AD4	Vina	QPLD	CS	CS rank	CR rank	RMSD
54	−53.4	−47.3	−45.4	−54.0	−53.7	−50.8	8	7	
55	−51.6	−43.1	−47.1	−51.0	−39.7	−46.5	22	22	
56	−58.5	−41.5	41.4	−32.2	−62.9	−30.7	84	39	
57	−47.4	−41.0	−44.6	−48.5	−16.7	−39.7	47	46	
58	−57.6	−51.4	−37.9	−52.7	−16.7	−43.3	37	28	
59	−48.2	−44.9	−29.2	−47.7	−16.7	−37.3	52	53	
60	−56.5	−49.5	−42.8	−52.7	−39.5	−48.2	13	17	
61	−43.6	−49.3	−40.9	−53.1	−38.5	−45.1	27	33	
62	−43.5	−49.0	−39.4	−51.0	−56.2	−47.8	17	19	
63	−44.1	−49.6	−33.5	−52.3	−40.8	−44.0	32	31	
64	−49.9	−47.2	−34.4	−45.6	−39.2	−43.3	37	44	
65	−9.2	−16.7	−42.2	−24.7	−16.7	−21.9	100	93	
66	−50.7	−43.3	−37.9	−44.8	−40.7	−43.5	36	36	
67	−42.2	−43.0	−11.8	−35.1	−41.0	−34.6	65	55	
68	−49.9	−47.2	−33.2	−45.6	−39.2	−43.0	40	45	
69	−40.9	−51.6	−15.8	−40.6	−41.3	−38.0	49	47	
70	−52.0	−39.0	−49.5	−52.3	−39.7	−46.5	23	23	
71	−43.9	−44.2	−39.2	−49.8	−41.8	−43.8	34	25	
72	−55.4	−50.0	−36.9	−48.1	−39.2	−45.9	26	32	
73	−34.0	−36.6	−36.3	−28.0	−40.5	−35.1	62	65	
74	−38.1	−39.2	−26.7	−21.8	−39.0	−33.0	74	75	
75	−24.8	−33.3	−35.8	−29.3	−37.8	−32.2	79	87	
76	−39.2	−35.9	−28.4	−27.6	−41.2	−34.4	67	60	
77	−32.5	−16.7	−24.5	−25.9	−38.9	−27.7	91	97	
78	−34.9	−37.1	−20.3	−14.2	−36.4	−28.6	90	96	
79	−21.7	−37.2	−29.5	−28.5	−16.7	−26.7	96	90	
80	−41.4	−33.6	−31.1	−26.4	−16.7	−29.8	86	79	
81	−37.0	−24.8	−26.0	−24.3	−39.7	−30.4	85	89	
82	−40.9	−38.5	−23.0	−26.4	−41.8	−34.1	68	57	
83	−27.3	−39.1	−27.2	−20.5	−39.7	−30.8	83	86	
84	−36.2	−38.1	−25.6	−28.0	−40.9	−33.8	72	66	
85	−14.6	−36.1	−29.9	−27.6	−40.4	−29.7	88	76	
86	−16.7	−16.7	−13.2	−16.7	−33.6	−19.4	102	102	
87	−34.6	−27.4	−29.7	−30.1	−39.4	−32.2	77	83	
88	−14.5	−35.5	−24.8	−18.4	−38.5	−26.4	97	98	
89	−40.0	−36.9	−23.6	−20.9	−41.6	−32.6	76	68	
90	−38.7	−22.0	−42.8	−36.4	−41.8	−36.3	54	51	
91	−35.8	−35.1	−39.0	−30.1	−39.7	−35.9	56	63	
92	−33.7	−32.3	−35.8	−33.5	−37.6	−34.6	66	82	
93	−33.3	−34.5	−36.9	−29.7	−34.9	−33.8	70	79	
94	−32.7	−36.3	−39.2	−29.7	−38.4	−35.3	61	69	
95	−33.4	−34.0	−19.5	−20.5	−41.7	−29.8	87	85	
96	−42.0	−37.0	−20.5	−15.5	−41.8	−31.3	81	66	
97	−32.5	−28.0	−31.1	−28.9	−16.7	−27.5	93	94	
98	−40.3	−35.2	−22.3	−23.0	−40.2	−32.2	78	77	
99	−36.0	−30.8	−37.0	−26.4	−39.4	−33.9	69	74	
100	−38.7	−37.5	−38.0	−22.6	−39.6	−35.3	60	64	
101	−22.1	−35.5	−36.3	−25.1	−16.7	−27.1	94	91	
102	−45.0	−16.7	4.0	−11.7	−40.8	−22.0	99	87	

The docking scores are in kJ/mol

CS consensus score, *CR* consensus rank, *RMSD* root-mean-quared deviation from the crystal structures in Å [27]

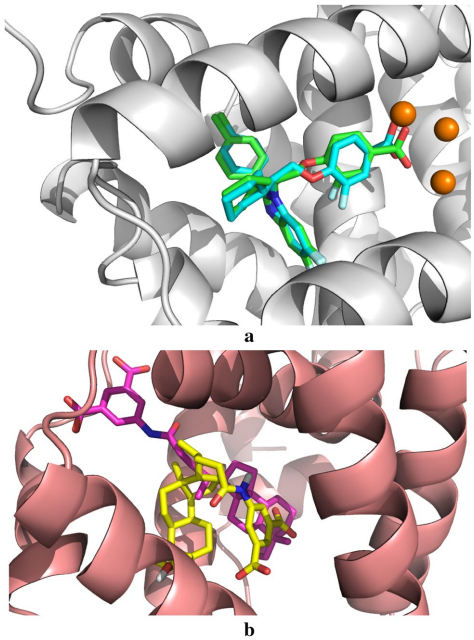


Fig. 2 The docked poses for (**a**) compound **28** (*cyan*), which gave the lowest RMSD (1.1 Å) among our results, compared to the crystal structure (protein in *white*, ligand in *green*, water molecules in *orange*) and (**b**) for ligand **34** (*yellow*), which gave the highest RMSD (9.6 Å) among our submissions, compared to the crystal structure (protein in *salmon*, ligand in *magenta*)

to obtain relative affinities of all the ligands and also to check the convergence with some thermodynamic cycles, as is shown in Fig. 1. The transformations were selected to minimise the difference between the ligands and they involve changes ranging from single-atom transformations (e.g. H→F/Cl/Br) to the introduction of a –CO–morpholine group. Five of the transformations involved a change in the net charge of the ligand and therefore required correction terms when simulated under periodic boundary-conditions with Ewald summation. We have therefore implemented the procedure suggested by Rocklin et al. [24] in connection with FEP free energies calculated with the AMBER software.

The results of the various FEP calculations are presented in Table 2. Compared to the experimental results [27], we obtained mean absolute deviation (MAD) of 7.5 ± 0.4 kJ/mol. This is slightly worse than in previous retrospective studies (4–6 kJ/mol) [9–12], but better than in the previous D3R Grand Challenge 2015 (4–16 kJ/mol) [15, 16]. The MAD was somewhat lower for set 1 (6.4 ± 0.5 kJ/mol) than for set 2 (8.6 ± 0.5 kJ/mol). The correlation between the calculated and experimental results was low, $R^2 = 0.08 \pm 0.02$. It was similar for the two sets, as can also be seen in Fig. 3. The τ_r was also poor, 0.05 ± 0.11, but it improved if relative affinities (both computed and experimental) were considered only if they were significantly different from zero at the 95% level ($\tau_{95} = 0.29 \pm 0.04$) [67]. This reflects that there are many experimental relative affinities with a small magnitude and therefore an uncertain sign (cf. Table 2). It seems more reasonable to exclude these in the calculations of τ.

As mentioned above, five of the studied transformations involved a change in the net charge of the ligand and for these we employed the semi-analytic correction suggested by Rocklin et al. [24]. As can be seen from Table 2, this correction amounted to 7–8 kJ/mol in all cases, with a positive sign if the starting ligand was charged and a negative sign if the final ligand was charged (the net charge of the ligands was either 0 or −1). The individual terms are shown in Table S3. The charge-correction calculations took only ~5 min/snapshot and can easily be automatised. For the four transformations with experimental data available, the charge correction always led to a reduced error and in two of the cases, it also corrected the sign of the result. Thus, it improved all quality measures (without the correction MAD = 8.1 ± 0.4 kJ/mol, $R^2 = 0.03 \pm 0.01$ and $\tau_r = -0.05 \pm 0.10$). Thus, the charge correction seems to be reliable and significantly improves the results. Excluding the four charge perturbations from the evaluation gave slightly better quality measures (MAD = 7.1 ± 0.4 kJ/mol, $R^2 = 0.11 \pm 0.02$ and $\tau_r = 0.06 \pm 0.10$) than if they were included, but the improvements are small and none of them is statistically significant.

Still, the largest deviation was observed for the **41→12** transformation in set 2 (22 kJ/mol), which involves the transformation of a benzoate group to the corresponding methyl ester, i.e., a charge perturbation. On the other hand, the other three charge transformations had smaller errors, 4–13 kJ/mol, and there was no correlation between the sign of the charge correction and the error. The other four transformations with an error larger than 15 kJ/mol involved the largest perturbation (**102→91**, i.e. –CO-morpholine→H), the introduction of a $–OCF_3$ group (**17→45**) and two simple H→Cl transformations (**77→12** and **81→85**). From this, it is hard to suggest a general explanation of the poor results for many of the transformations.

The precision of the calculated affinities is also given in Table 2. It can be seen that it was small for all transformations, 0.1–0.7 kJ/mol. The charge correction added an extra term with an uncertainty of 0.2–0.3 kJ/mol, so these transformations always gave the higher uncertainties (0.5–0.7 kJ/mol, compared to 0.1–0.5 kJ/mol for the other transformations). However, the charge perturbations gave a high uncertainty already without the charge corrections (0.4–0.4 kJ/

Table 2 Calculated (with and without charge correction) and experimental [27] relative binding free energies (kJ/mol) for the two FEP sets

Perturbation	$\Delta\Delta G_{bind}$		$\Delta\Delta G_{CC}$		$\Delta\Delta G_{bind,corr}$		$\Delta\Delta G_{exp}$
FEP Set 1							
17→45	−4.6	±0.2			−4.6	±0.2	10.6
17→49	8.5	±0.2			8.5	±0.2	14.3
17→91	16.7	±0.3			16.7	±0.3	10.7
45→91	11.4	±0.4			11.4	±0.4	0.1
46→49	−0.3	±0.1			−0.3	±0.1	1.4
47→91	−3.3	±0.3			−3.3	±0.3	1.0
48→91	6.0	±0.4			6.0	±0.4	−3.6
49→91	1.5	±0.3			1.5	±0.3	−3.6
93→91	5.7	±0.1			5.7	±0.1	−1.3
95→91	−1.0	±0.3			−1.0	±0.3	−0.2
96→46	2.5	±0.3			2.5	±0.3	0.2
96→91	7.4	±0.3			7.4	±0.3	−2.0
96→98	2.5	±0.2			2.5	±0.2	−4.4
98→46	4.3	±0.2			4.3	±0.2	4.6
98→91	9.8	±0.3			9.8	±0.3	2.4
99→91	2.4	±0.2			2.4	±0.2	−3.6
100→91	3.4	±0.4			3.4	±0.4	1.3
101→91	−4.5	±0.6	8.4	±0.3	3.9	±0.7	0.2
102→91	17.4	±0.4			17.4	±0.4	0.0
FEP Set 2							
10→73	−17.4	±0.5	6.9	±0.2	−10.5	±0.6	2.0
10→79	11.7	±0.5			11.7	±0.5	−0.9
12→76	6.9	±0.2			6.9	±0.2	19.4
38→10	7.1	±0.4	−7.9	±0.3	−0.8	±0.5	−8.5
41→12	7.2	±0.4	−6.9	±0.2	0.4	±0.5	−21.9
41→38	5.2	±0.2			5.2	±0.2	0.0
41→M1	2.1	±0.2			2.1	±0.2	
73→75	11.5	±0.2			11.5	±0.2	6.4
74→76	5.4	±0.2			5.4	±0.2	12.2
76→10	2.3	±0.2			2.3	±0.2	−5.9
77→12	11.2	±0.1			11.2	±0.1	−4.3
77→82	5.5	±0.2			5.5	±0.2	−1.0
78→12	−0.8	±0.2			−0.8	±0.2	2.1
81→85	11.5	±0.3			11.5	±0.3	−6.5
82→84	8.8	±0.2			8.8	±0.2	9.5
83→12	0.8	±0.2			0.8	±0.2	−5.1
84→76	3.4	±0.1			3.4	±0.1	6.5
85→76	3.2	±0.1			3.2	±0.1	14.6
88→76	5.7	±0.2			5.7	±0.2	12.8
88→85	−2.0	±0.3			−2.0	±0.3	−1.8
89→76	5.1	±0.2			5.1	±0.2	11.9
M1→84	9.7	±0.4	−6.8	±0.2	2.9	±0.5	

mol), reflecting that a change of the net charge of the ligand gives rise to large fluctuations of the electrostatic interaction with the surrounding protein. Still, it is clear that the rather poor results (e.g. MAD = 7.5 kJ/mol) are not caused by a too low precision (0.1–0.7 kJ/mol).

Likewise, there is no indication of any poor overlap in any of the studied transformations. On the contrary, the seven overlap measures listed in Table S2 all indicate proper overlap throughout the transformations. In fact, we first run some of the transformations with only 13 λ values, but the overlap

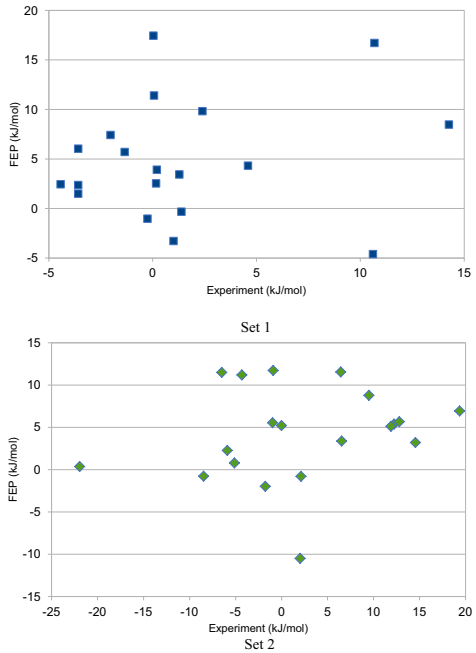


Fig. 3 Comparison between the experimental [27] and calculated binding free energies for the two FEP sets

measures sometimes indicated poor overlap. Therefore, we decided to use 25 λ values for all transformations.

On the other hand, the thermodynamic cycles indicate an appreciably poorer convergence of the results, as can be seen in Table 3. Of the nine studied cycles, only two gave a vanishing result, within the statistical precision, both in FEP set 2, one of which involves two charge-perturbation steps and the extra **M1** ligand (**76→12→41→M1→84→76**, 1.2 ± 0.7 kJ/mol; also **76→12→77→82→84→76**, −0.4 ± 0.4 kJ/mol). The other six cycles gave larger hystereses, 4–10 kJ/mol. The one with the largest hysteresis (**91→17→45→91**) involves the two perturbations (**45→91** and **17→45**) for which BAR and MBAR gave results that differ significantly (by 1.5 and 4.6 kJ/mol; cf. $\Delta\Delta G_{\mathrm{BAR}}$ in Table S2), whereas for all the other transformations, the difference was less than 1.2 kJ/mol (0.5 kJ/mol on average). They involve the introduction of $-OCF_3$ and –COOEt groups. Large cycle-closure errors indicate that sampling of the phase space has been incomplete. This may be caused by a change of the binding mode of the ligands. However, we have not been able to identify such problems by overlaying the structures. The problems could perhaps have been

Table 3 Thermodynamic cycles and the cycle hysteresis (kJ/mol)

Cycle	Hysteresis
91→49→17→91	6.7 ± 0.4
91→17→45→91	−9.9 ± 0.6
91→96→46→49→91	−3.7 ± 0.5
96→98→46→96	4.2 ± 0.4
91→96→98→91	4.9 ± 0.5
76→12→41→M1→84→76	1.2 ± 0.7
76→12→77→82→84→76	−0.4 ± 0.4
76→12→41→38→10→76	−5.1 ± 0.8
76→85→88→76	4.5 ± 0.4

solved by longer simulations or enhanced-sampling techniques. Alternatively, several independent perturbations could have been run, which often give a better estimate of the true uncertainty and a more effective sampling of the phase space [13, 74, 75]. In fact, test calculations indicated that $\Delta\Delta G_{\mathrm{bind}}$ from independent repeats varies by ~2 kJ/mol.

In the GC2 evaluation, the relative binding affinities were recalculated to absolute affinities, by employing **10** and **17** as reference ligands for sets 1 and 2, respectively. This makes the results dependent on the selected reference ligand (ligands **76** and **91** would have been more natural, based on our perturbation networks, shown in Fig. 1, whereas ligand **10** is very peripheral) and make the uncertainties more varying, as they depend on the number of perturbations needed to reach the various ligands from the selected reference. Still, this is necessary to enable a comparison between the various methods.

In the evaluation of the various submissions (22 for both FEP sets, although only 18 and 19 involved all ligands for set 1 and 2, respectively), our results gave $\tau = 0.02 \pm 0.22$, $\rho = 0.12 \pm 0.27$, $R = 0.34 \pm 0.27$ and RMSD = 6.3 ± 1.3 kJ/mol for set 1 and $\tau = 0.48 \pm 0.14$, $\rho = 0.66 \pm 0.14$, $R = 0.58 \pm 0.13$ and RMSD = 6.3 ± 0.8 kJ/mol for set 2. R for set 1 was the second best among all submissions, whereas most of the other entries ranked number five, except τ and ρ for set 2 (12–16). However, our method gave relatively accurate results for both sets and also comparable results for all measures, whereas most other methods gave more varying results. Therefore, our method was among the four submissions that gave the best results for both FEP sets. Two of the other top submissions also employed FEP, using the Schrödinger software and the OPLS3 force field (submissions pyxiv and x2j7p by Cournia group and submissions ck8kc and 81n55 by an anonymous group). Both gave the same average RMSD as our submission, 6 ± 1 kJ/mol. The third submission (3idpo and rvm67 by Camacho group), used instead the "quasi-exact" scoring approach, which actually gave the lowest RMSD for set 1, 4.9 kJ/mol, but worse average τ, ρ and R results than our submission. None of the

four quality estimates showed any statistically significant differences for any of the two FEP sets between our results and those of the other three top submissions. FEP calculations by the Michel group also gave low RMSD, but they had problems with the charge perturbations and the best results were obtained when those perturbations were excluded. The FEP calculations with the Schrödinger software employed only neutralised ligands. Other approaches, including MM/GBSA, MM/PBSA, multi-site lambda dynamics and also one set of FEP calculations gave clearly worse results.

Our FEP results can also be compared to those obtained with the consensus score (CS) from our docking calculations. To this end, we took the difference of the CS results for the two ligands involved in the same perturbations studied by FEP (Table 2). Interestingly, CS gave results of nearly the same quality as FEP: the MAD was slightly lower for FEP set 1 (5.5 ± 0.8 kJ/mol compared to 6.4 ± 0.5 kJ/mol), but slightly higher for set 2 (10.0 ± 0.9 kJ/mol, compared to 8.6 ± 0.5 kJ/mol; standard errors for CS were estimated from the standard deviation over the $\Delta\Delta G_{\text{bind}}$ results for each of the five scoring methods and it was much higher than for FEP, 1–9 kJ/mol). On the other hand, the correlation was worse for both sets, $R^2 = 0.04 \pm 0.04$ and -0.46 ± 0.10 (i.e. an anticorrelation), compared to 0.09 ± 0.02 and 0.08 ± 0.02. τ_r was slightly better for set 1, but appreciably worse for set 2, 0.16 ± 0.21 and -0.47 ± 0.19, compared to 0.05 ± 0.11 and 0.05 ± 0.12. The poor τ results, compared to those calculated for all 102 ligands (0.26 ± 0.06), indicates that the binding affinities in the FEP sets were harder to estimate than the those of the other ligands.

Conclusions

In this investigation, we have studied the binding of 102 ligands to FXR from the blind-prediction D3R Grand Challenge 2016 with five different docking and scoring methods. Considering that we only provided a single pose for each ligand, the results were decent, in the middle among the GC2 submissions, and comparable to some FEP results. The scoring gave fairly good results with a τ of 0.26 ± 0.06 and a ρ of 0.41 ± 0.08, especially considering that only one protein structure was used for all ligands. Better results may perhaps have been obtained with more relevant crystal structures or considering more flexibility of the binding site, fixing parts of the ligand, demanding certain protein–ligand interactions to be fulfilled for the docking programs, using more than one binding pose for scoring or using even higher exhaustiveness settings.

Moreover, we have employed a FEP protocol to calculate relative binding free energies for the 33 ligands in the FEP set. In particular, we have implemented and benchmarked the approach of Rocklin et al. [24] to correct for artefact caused by the periodic simulations with Ewald summation for transformations that changed the net charge of the ligand. The accuracy is slightly worse than in retrospective large-scale tests of FEP methods [9–12] (MAD = 7.5 kJ/mol, $R^2 = 0.1$ and $\tau_{r,95} = 0.3$), but better than in the D3R Grand Challenge 2015 [15, 16]. The charge corrections are significant (7–8 kJ/mol) and always improve the results. The precision of the estimated binding affinities is good (0.1–0.7 kJ/mol) and our measures indicate that the overlap throughout the transformations is excellent, owing to the use of 25 λ values. However, the thermodynamic cycles indicate that the sampling in several cases has been unsatisfactory. This could have been resolved by more simulations (although the time was limited). Moreover, it is possible that we have employed incorrect structures or that the binding mode changes for the various ligands (only three crystal structures are available for the studied ligands), which may explain the rather poor results.

Interestingly, FEP calculations with the Schrödinger FEP software and the latest force field OPLS3 [12] did not give any significantly better results, although they involved longer simulations (5 ns), enhanced-sampling methods and automatic mapping of the ligands. The reason for this may be that they did not employ any charge corrections, but instead supposed that all ligands were neutral when binding. The prime conclusion of this prospective study is that the charge corrections are large (7–8 kJ/mol) and significantly improve the results. The correction employed in this investigation [24] is easy to implement and does not increase the computational load significantly.

Acknowledgements This investigation has been supported by grants from the Swedish research council (project 2014–5540), the Knut and Alice Wallenberg Foundation (KAW 2013.0022) and the Estonian Ministry of Education and Research (A.T.G.-S. Grant IUT34-14). 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. Jorgensen WL (2009) Acc Chem Res 42:724–733
2. Gohlke H, Klebe G (2002) Angew Chem Int Ed 41:2644–2676
3. Wereszczynski J, McCammon JA (2012) Quart Rev Biophys 45:1–25
4. Hansen N, van Gunsteren WF (2014) J Chem Theory Comput 10:2632–2647
5. Zwanzig RW (1954) J Chem Phys 22:1420–1426
6. Kirkwood JG (1935) J Chem Phys 3:300–313

7. Bennett CH (1976) J Comput Phys 22:245–268
8. Shirts MR, Chodera JD (2008) J Chem Phys 129:124105
9. Christ C, Fox TJ (2014) Chem Inf Model 54:108–120
10. Mikulskis P, Genheden S, Ryde UJ (2014) Chem Inf Model 54:2794–2806
11. Wang L, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, Romero DL, Masse C, Knight JL, Steinbrecher T, Beuming T, Damm W, Harder E, Sherman W, Brewer M, Wester R, Murcko M, Frye L, Farid R, Lin T, Mobley DL, Jorgensen WL, Berne BJ, Friesner RA, Abel RJ (2015) Am Chem Soc 137:2695–2703
12. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA (2016) J Chem Theory Comput 12:281–296
13. Muddana HS, Fenley AT, Mobley DL, Gilson MK (2014) J Comput Aided Mol Des 28:305–317
14. Mikulskis P, Cioloboc D, Andrejic M, Khare S, Brorsson J, Genheden S, Mata RA, Söderhjelm P, Ryde U (2014) J Comp-Aided Mol Design 28:375–400
15. Gathiaka S, Liu S, Chiu M, Yang H, Stuckey JA, Kang YN, Delproposto J, Kubish G, Dunbar JB, Carlson HA, Burley SK, Walters WP, Amaro RE, Feher VA, Gilson MK (2016) J Comput Aided Mol Des 30:651–668
16. Misini Ignjatovic M, Caldararu O, Dong G, Muñoz-Gutierrez C, Adasme-Carreño F, Ryde U (2016) J Comp-Aided Mol Design 30:707–730
17. Kastenholz MA, Hünenberger PH (2006) J Chem Phys 124:224501
18. Simonson T, Roux B (2016) Mol Sim 42:1090–1101
19. Genheden S, Nilsson I, Ryde U (2011) J Chem Inf Model 51:947–958
20. Genheden S, Ryde U (2012) J Chem Theory Comput 8:1449–1458
21. Gl Warren, Patel S (2007) J Chem Phys 127:064509
22. Reif MM, Hünenberger PH (2011) J Chem Phys 134:144103
23. Ekimoto T, Matubayasi N, Ikeguchi M (2015) J Chem Theory Comput 11:215–223
24. Rocklin GJ, Mobley DL, Dill KA, Hünenberger PH (2013) J Chem Phys 139:184103
25. Reif MM, Oostenbrink C (2014) J Comput Chem 35:227–243
26. Pellicciari R, Costantino G, Fiorucci S (2005) J Med Chem, 48:5383
27. FXR experimental data for the D3R Grand Challenge 2. https://drugdesigndata.org/about/grand-challenge/413
28. Ryan KK, Tremaroli V, Clemmensen C, Kovatcheva-Dachary P, Myronovych A, Karns R, Wilson-Pérez HE, Sandoval DA, Kohli R, Bäckhed F, Seeley RJ (2014) Nature 509:183–188
29. Richter HGF, Benson GM, Bleicher KH, Blum D, Chaput E, Clemann N, Feng S, Gardes C, Grether U, Hartman P, Kuhn B, Martin RE, Plancher JM, Rudolph MG, Schuler F, Taylor S (2011) Bioorg Med Chem Lett 21:1134–1140
30. Case DA, Berryman JT, Betz RM, Cerutti DS, Cheatham TE, Darden III, T.A., Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand S, Li P, Luchko T, Luo R, Madej B, Merz KM, Monard G, Needham P, Nguyen H, Nguyen HT, Omelyan I, Onufriev A, Roe DR, Roitberg A, Salomon-Ferrer R, Simmerling CL, Smith W, Swails J, Walker RC, Wang J, Wolf RM, Wu X (2015) AMBER. York DM, Kollman PA (eds). University of California, San Francisco
31. Maestro version 10.2 (2015) Maestro version 10.2. Schrödinger LLC, New York, NY
32. http://signe.teokem.lu.se/~ulf/Methods/ekvilibr.html, point 10
33. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T (2004) J Chem Phys 120:9665–9678
34. Cho AT, Guallar V, Berne BJ, Friesner R (2005) Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) aproach. J Comput Chem 26:915–931
35. Small molecule package (2016) Small molecule package. Schrödinger LLC, New York
36. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJJ (1998) Comput Chem 19:1639–1662
37. Trott O, Olson AJ (2010) J Comput Chem 31:455–461
38. LigPrep version 3.4 (2015) LigPrep version 3.4. Schrödinger LLC, New York, NY
39. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) J Comput Chem 16:2785–2791
40. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch ZE, Hutchinson GR (2012) J Chem Inform 4:17
41. Rappe AK, Casewit CJ, Colwell KS, Goddard WA, Skiff WM (1992) J Am Chem Soc 114:10024–10035
42. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) J Chem Theory Comput 11:3696–3713
43. Wang JM, Wolf RM, Caldwell KW, Kollman PA, Case DA (2004) J Comput Chem 25:1157–1174
44. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902–3909
45. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56:2257
46. Besler BH, Merz KM, Kollman PA (1990) J Comput Chem 11:431–439
47. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision A02. Gaussian Inc, Wallingford CT
48. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) J Phys Chem 97:10269–10280
49. Seminario JM (1996) Int J Quant Chem 60:1271
50. Nilsson K, Lecerof D, Sigfridsson E, Ryde U (2003) Acta Crystallogr D 59:274–289
51. Becke AD (1988) Phys Rev A 38:3098–3100
52. Lee CT, Yang WT, Parr RG (1988) Phys Rev B, 37:785
53. Weigend F, Ahlrichs R (2005) Phys Chem Chem Phys 7:3297–3305
54. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
55. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA (2009) Small molecule hydration free energies in explicit solvent: an extensive test of fixed-charge atomistic simulations. J Chem Theor Comput 5:350–358
56. Wong V, Case DA (2008) Evaluating Rotational Diffusion from Protein MD Simulations. J Phys Chem B 112:6013–6024
57. Kaus JW, Pierce LT, Walker RC, McCammon JA (2013) J Chem Theory Comput 9:4131–4139
58. Steinbrecher T, Mobley DL, Case DA (2007) J Chem Phys 127:214108
59. Steinbrecher T, Joung I, Case DA (2011) J Comp Chem 32:3253–3263

60. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
61. Wu X, Brooks BR (2003) Self-guided langevin dynamics simulation method. Chem Phys Lett 381:512–518
62. Berendsen HJC., Postma JPM., Van Gunsteren WF, Dinola A, Haak JR (1984) Molecular Dynamics with Coupling to an External Bath. J Chem Phys 81:3684–3690
63. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. J Chem Phys 98:10089–10092
64. Tembe BL, McCammon JA (1984) Comp Chem 8:281–283
65. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Proc Natl Acad Sci USA 98:10037–10041
66. Sitkoff D, Sharp KA, Honig B (1994) J Phys Chem 98:1978–1988
67. Mikulskis P, Genheden S, Rydberg P, Sandberg L, Olsen L, Ryde U (2012) J Comput-Aided Mol Design 26:527–541
68. Genheden S, Ryde UJ (2010) Comput Chem 31:837–846
69. Brown SP, Muchmore SW, Hajduk PJ (2009) Drug Discov Today 14:420–427
70. Bhattacharyya A (1943) Bull Cal Math Soc 35:99–109
71. Wu D, Kofke DAJ (2005) Chem Phys 123:1–10
72. Rod TH, Ryde U (2005) Phys Rev Lett 94:1–4
73. Zhong H, Kirschner KN, Lee M, Bowen JP (2008) Binding free energy calculation for duocarmycin/DNA complex based on the QPLD-derived partial charge model. Bioorg Med Chem Lett 18:542–545
74. Genheden S, Ryde U (2010) J Comput Chem, 31:837–846
75. Genheden S, Ryde U (2011) J Comput Chem 32:187–195

Paper VI

Maria Luisa Vertaramo, Olof Stenström, Majda Misini Ignjatović,
Octav Caldararu, Martin A. Olsson,
Francesco Manzoni, Hakon Leffler, Esko Oksanen, Derek T. Logan,
Ulf J. Nilsson, Ulf Ryde, and Mikael Akke, Manuscript, **2018**

Detailed characterization of the binding of diastereomeric ligands to galectin-3

Maria Luisa Vertaramo[‡,1], Olof Stenström[‡,2], Octav Caldararu[‡,3], Majda Misini Ignjatovic[‡,3], Martin A. Olsson[‡,3], Francesco Manzoni[‡,3,4,†], Hakon Leffler[5], Esko Oksanen[6], Derek T. Logan[4], Ulf J. Nilsson[1], Ulf Ryde[3], and Mikael Akke*[,2]

[1] Centre for Analysis and Synthesis, Department of Chemistry, Lund University; [2] Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University; [3]Theoretical Chemistry, Department of Chemistry, Lund University; [4]Biochemistry and Structural Biology, Center for Molecular Protein Science, Department of Chemistry; [5]Microbiology, Immunology, and Glycobiology, Department of Experimental Medicine, Lund University; [6]European Spallation Source ESS ERIC

ABSTRACT: Understanding the driving forces underlying molecular recognition is of fundamental importance in chemistry and biology. Here we pinpoint the molecular determinants behind differences in affinity of two diastereomeric ligands binding to the carbohydrate recognition domain of galectin-3. Because the two diastereomers have essentially equal chemical potential in the free state, the free states of the protein and ligand cancel in the comparative analysis of binding thermodynamics, which makes it possible to explain the difference in binding affinity based solely on the statistical-mechanical properties of the two ligand–protein complexes. We characterized the binding of the two ligand R- and S-stereoisomers to galectin-3 by a combination of isothermal titration calorimetry, X-ray crystallography, NMR relaxation, and molecular dynamics simulations followed by calculations of conformational entropy and grid inhomogeneous solvation theory. The results show that while the two diastereomers have nearly equal free energies of binding, $\Delta\Delta G°$(R–S) = –1.9 ± 0.1 kJ/mol, they exhibit greater differences in their overall enthalpies, $\Delta\Delta H°$(R–S) = –5 ± 1 kJ/mol, and entropies of binding, $-T\,\Delta\Delta S°$(R–S) = 3 ± 1 kJ/mol. The crystal structures of the two complexes reveal that the deepest buried segment of the ligands have indistinguishable structure and interactions with the protein, and a key hydroxyl group at the stereo-center of the ligand is coordinated identically in the two complexes, while the more solvent-exposed segments of the ligands differ in their bound conformations. Notably, crystallography, NMR and molecular dynamics simulations indicate that the protein in complex with the S-stereoisomer has considerably greater conformational entropy than in the other complex. On the other hand, the difference in solvation entropy, associated with key water sites around the bound ligand, is quite small. These results illustrates how the binding of similar ligands involve a subtle interplay between protein conformational fluctuations and variations in water coordination at the binding site, which points to both opportunities and challenges in rational drug design.

INTRODUCTION

Molecular recognition is fundamental to biology in that it governs signaling within and between cells, with prominent examples provided by the immune system, hormonal control of distant organs in higher organisms, and specificity of enzyme reactions. Modern medicine is to a large extent based on the possibility to interfere with and control molecular recognition by the design of synthetic ligands or effectors that bind to a specific protein in a given signaling pathway. Drug design aims to generate such protein ligands that have high affinity and specificity for the target. Despite enormous resources contributed by industry and academia over the last several decades, rational structure-based design of ligands by computational approaches remains extremely challenging. One reason is that the free energy of binding is in most cases a small difference between large numbers arising from the different interactions between the protein, ligand, other solutes, and solvent molecules. In addition, the energy terms are strongly dependent on the detailed molecular conformations, due to their sharp dependence on interatomic distances and orientations. Furthermore, entropic contributions can be significant because proteins have many degrees of freedom, are generally flexible, and consequently populate a wide range of conformations. Recent work has indeed highlighted the role of protein conformational entropy in ligand binding,[1–8] as well as the highly heterogeneous response of water molecules around binding sites.[9–12]

We have identified the carbohydrate recognition domain (CRD) of galectin-3 (denoted galectin-3C) as an interesting system for investigating the role of conformational entropy[4,5] and

solvation in ligand binding.[13] Galectin-3 has a relatively solvent-accessible binding site placed in a shallow groove across one of the two β-sheets, with water molecules forming an integral part of the binding site by bridging between the ligand and protein.[13] Galectin-3 is a member of the galectin family of mammalian lectins, defined by the CRD with its conserved sequence motif that confers affinity for β-galactoside containing glycans.[14,15] Galectins play important roles in cell growth, cell differentiation, cell cycle regulation, signaling, and apoptosis, which target them for pharmaceutical intervention to treat inflammation and cancer, [16,17] with specific examples reported for galectin-3.[18–20]

Here, we report a comparative analysis of galectin-3C in complex with two diastereomeric ligands. The advantage of this approach is that the differences in binding thermodynamics are dominated by the properties of the two ligand–protein complexes, while the unbound diastereomers have nearly identical chemical potential in the unbound state and thus cancel in the comparative analysis. We used a combination of experimental and computational approaches including isothermal titration calorimetry (ITC), competitive fluorescence polarization assay, X-ray crystallography, NMR spectroscopy including ^{15}N backbone and ^{2}H side-chain methyl relaxation, and molecular dynamics (MD) simulations followed by conformational entropy and grid inhomogeneous solvation theory (GIST) calculations.

MATERIALS AND METHODS

Ligand synthesis. The two diastereomeric compounds (2R)- and (2S)-2-hydroxy-3-(4-(3-fluorophenyl)-1H-1,2,3-triazol-1-yl)-propyl 2,4,6-tri-O-acetyl-3-deoxy-3-(4-(3-fluorophenyl)-1H-1,2,3-triazol-1-yl)-1-thio-β-D–galactopyranoside (denoted ligands R and S, respectively) were synthesized from triisopropylsilyl 2,4,6-tri-*O*-acetyl-3-azido-3-deoxy-1-thio-β-D-galactopyranoside[21] and R- and S-glycidyl nosylate. Reaction conditions, physical data, and purity data are given in SI.

Protein expression and purification. Galectin-3C was expressed and purified by the Lund Protein Production Platform (LP3) at Lund University following published protocols,[4,5] yielding a protein stock solution of 9.2 mg/ml in ME-PBS buffer (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.3, 2 mM ethylenediaminetetraacetic acid (EDTA), 4 mM β-mercaptoethanol), and 150 mM lactose. The protein stock solution was stored at 278 K.

Isothermal titration calorimetry. Galectin-3C samples were prepared by extensive dialysis against 5 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES) buffer to remove all lactose, followed by centrifugation at 14,000 rpm to remove any aggregates. Both ligands were dissolved in stock solutions of dimethylsulfoxide (DMSO) to prepare stock solutions 20.7 mM and 20.3 mM for R and S, respectively, and sonicated immediately prior to experiments. Isothermal titration calorimetry (ITC) experiments were performed on MicroCal iTC200 and MicroCal PEAQ–ITC instruments (Malvern) at a temperature of 301 K by titrating the protein at a concentration of 0.22 mM into the cell containing the ligand at a concentration of 0.02 mM. The DMSO concentrations in the cell and the syringe were carefully matched to minimize the heat of dilution. Five replicate experiments were performed for each complex. Peak integration was done using NITPIC.[22] A single-site binding model was fitted simultaneously to the 5 titrations curves to yield the binding enthalpy (ΔH), fraction of binding-competent protein (n), and dissociation constant (K_d), using in-house routines implemented in Matlab with Monte Carlo error estimation.[23] The heat released or absorbed during the ith injection is given by[24]

$$\Delta Q(i) = Q(i) - Q(i-1) + (V_i/V_0)[Q(i) - Q(i-1)]/2 + Q_{off}$$

where V_i is the volume of the ith injection, V_0 is the cell volume, Q_{off} is an offset parameter that accounts for heats of mixing, and $Q(i)$ is the heat function following the ith injection:

$$Q = (\Delta V_0/2)[nM_t + X_t + K_d - \{(nM_t + X_t + K_d)^2 - 4nM_tX_t\}^{1/2}]$$

where M_t and X_t are the total concentration of the protein and ligand, respectively, in the cell at any given point of the titration. The free energy and entropy of binding were subsequently determined using the relationships $\Delta G^\circ = RT\ln(K_d)$ and $-T\Delta S^\circ = \Delta G^\circ - \Delta H^\circ$.

Competitive fluorescence polarization experiments. The binding affinity between galectin-3C and each ligand was determined using competitive fluorescence polarization experiments described previously,[20] using the fluorescent probe 3,3′-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1H-1,2,3-triazol-1-yl]-3′-(3,5-dimethoxybenzamido)- 1,1′-sulfanediyl-di-β-D-galactopyranoside. [25]

X-ray crystallography. Small crystals of lactose-bound galectin-3C were grown with the hanging drop method in Nextal plates and with the following reservoir condition: 28% (w/v) PEG 4000, Tris-HCl pH 7.5, 0.4 M NaSCN, 15 mM β-mercaptoethanol. The drop volume was 2 μl and the protein solution:reservoir ratio was varied between 0.5:1, 1:1, and 2:1. Small crystals were then moved to drops containing the same reservoir with the addition of 10 mM of the ligand (R or S), from a 100 mM stock solution in neat DMSO. Soaking lasted for 7 hours for R and 20 hours for S. Before data collection, crystals were placed for a couple of seconds in a drop containing 1 volume of PEG400 100% and 3 volumes of crystallization solution to prevent ice formation during cryo-cooling. Data were collected at beamline I911-3 of the MAX-II synchrotron, Lund, Sweden.[26] All data were integrated using XDS.[27] Diffraction data for R were collected in a single pass, while that for S involved two passes, one at low resolution with lower exposure time followed by one at high resolution, and subsequently scaled and merged with XSCALE.[27]

MTZ files were generated with Aimless.[28] Cross validation was based on 10% of the reflections. An initial structure solution was determined through rigid-body refinement in Refmac5[29] using as a starting model the lactose–galectin-3C structure[13] with lactose and water molecules removed and with the resolution limit set to 3.5 Å. The R–galectin-3C and S–galectin-3C complex structures were built manually using Chimera[30] and crystallographic restraints were obtained through phenix.eLBOW.[31] Restrained refinement was then performed using phenix.refine[32] using all remaining reflections. Manual rebuilding, including addition of water molecules, was done using Coot.[33]

Ensemble refinement of crystal structures. Ensemble refinement of the X-ray diffraction data was performed using the

module *phenix.ensemble_refinement* in the *Phenix* software suite.[34] The X-ray crystal structures of the S–galectin-3C and R–galectin-3C complexes from the previous section were used as starting structures. The crystallographic water molecules were kept and hydrogen atoms and missing atoms in the protein were added using the Leap module from the Amber 14 software.[52] Ligand restraints and coordinates were the same as those used in the original refinement.

The collective dynamics of the protein was described using a TLS model with a single group, which included both the protein and the ligand atoms. A model including two TLS groups was also tested — one for the ligand and one for the protein — but it gave worse results (R_{free} values of 20.6–20.8 compared to 19.0 for the single TLS model). The percentage of atoms included in the TLS-fitting (p_{TLS}) was optimized by testing five different values (0.5, 0.6, 0.7, 0.8 and 0.9) and choosing the one that yielded the lowest R_{free}, which was p_{TLS} = 0.7 for both protein–ligand complexes. An ensemble of structures was then generated by running MD simulations, in which the model was restrained by a time-averaged X-ray maximum-likelihood target function. The X-ray weight-coupled temperature bath offset was kept at the default value of 5 K. A 1.25 ps relaxation time of the time-averaged-restraints was used, resulting in 25 ps long MD simulations, with structures stored every 0.05 ps. Owing to the good resolution of both data sets, ordered solvent molecules were not updated during the simulation. All structures generated by ensemble refinement were kept, resulting in 500 different structures in each ensemble. Atomic fluctuations were calculated using the *cpptraj* module of Amber after removal of the water molecules.[35]

NMR sample preparation. The galectin-3C concentration was 0.32, 0.2 and 0.34 mM for the ^{15}N, ^{15}N/^{13}C, and ^{15}N/^{13}C/^{2}H samples, respectively. The ligands were dissolved in neat DMSO to a concentration of 8.2 mM for S and 35 mM for R. The protein–ligand complexes were prepared by titrating the ligand into the protein, while monitoring the ^{15}N heteronuclear single-quantum correlation (HSQC) spectra. The final DMSO content in the NMR sample was 4.3% for S and 1.2% for R.

NMR resonance assignments. Backbone chemical shift assignments were based on HNCACB[36] spectra and previous assignments for various galectin-3C complexes.[5] Methyl groups were assigned using CCH-TOCSY and HCCH-TOCSY experiments.[37,38] All spectra were processed using NMRPipe,[39] employing a processing protocol including a solvent filter, square cosine apodization, and zero filling to twice the number of points in all dimensions. All spectra were analyzed using the CCPNmr program suite.[40]

NMR relaxation experiments and data analysis. ^{15}N R_1, R_2, and {^{1}H}–^{15}N nuclear Overhauser effect (*NOE*) experiments targeting the backbone amides were performed at magnetic field strengths of 11.7, 14.1 and 21.1 T, and a temperature of 301 K. Spectral widths were 14–16 ppm and 28–30 ppm for ^{1}H and ^{15}N, respectively, covered by 1024 and 128 points. Relaxation decays were recorded with 10 relaxation delays ranging between 0–1 s for R_1 acquired at 11.7 and 14.1 T, 0–3 s for R_1 acquired at 21.1 T, and 0–0.2 s for R_2 (at all fields) with a 1.2 ms delay between refocusing pulses. The *NOE* was measured using a ^{1}H saturation time of 7 s and a recycle delay between experiments of 3 and 7 s for experiments acquired at 11.7 and 14.1 T, respectively, while the reference experiment was acquired using a recycle delay of 10 and 14 s at 11.7 and 14.1 T, respectively. *NOE* experiments performed at 21.1 T employed a ^{1}H saturation time of 6 s and a recycle delay between experiments of 2 s, while the reference experiment was acquired with a recycle delay of 14 s. Peak intensities were evaluated as partial peak volumes calculated over 3 × 5 points in the direct and indirect dimension, respectively. Mono-exponential functions were fitted to the R_1 and R_2 relaxation decays using the CCPNmr program suite and bootstrap error estimation. *NOE*s were calculated as the ratio of the peak intensities in the saturated and reference experiments, and the standard errors were determined by propagating the errors of intensities estimated from the baseplane noise.

Backbone ^{15}N CPMG relaxation dispersion experiments[41] were recorded at static magnetic field strengths of 11.7 and 14.1 T for S–galectin-3C and at 14.1 T for R–galectin-3C. Relaxation dispersion profiles were acquired using 20 refocusing frequencies ranging from 50–1000 Hz and a constant relaxation time of 40 ms, using a two-point approximation of the exponential decay.[42]

$R_1(D_z)$, $R(3D_z^2–2)$, $R_2(D_+)$ and $R(D_+D_z+D_zD_+)$ ^{2}H relaxation experiments[43] targeting the methyl groups were recorded at 11.7 and 14.1 T. Spectral widths were 16 and 20 ppm for ^{1}H and ^{13}C, respectively, covered by 1024 points in the ^{1}H dimension at both field strengths, and 70 and 84 points for ^{13}C at 11.7 and 14.1 T, respectively. The number of points recorded were 1024 for ^{1}H at both static magnetic field strengths. Relaxation decays were sampled by 9 points covering 0–0.1 s for $R_1(D_Z)$ and $R(3D_z^2–2)$, 0–20 ms for $R_2(D_+)$ and $R(D_+D_z+D_zD_+)$. The recycle delay was 1.8–2 s. Peak volumes were evaluated using the program suite PINT.[44] Mono-exponential functions were fitted to the relaxation decays using an in-house Matlab script with Monte Carlo error analysis.[23]

Model-free analysis of NMR relaxation data. Backbone amide model-free parameters were fitted using the program suite relax,[45–47] using a N–H bond length of 1.02 Å and a ^{15}N chemical shift anisotropy of –172 ppm. The backbone optimization was restricted to five different models defined by the parameter sets: {O^2}, {O^2, τ_e}, {O^2, R_{ex}}, {O^2, τ_e, R_{ex}}, or {O^2_f, O^2_s, τ_s}, where O^2, O^2_f, and O^2_s denote the order parameter with subscripts *f* and *s* indicating that the order parameter can be resolved into amplitudes of fluctuation taking place on separate time scales (fast and slow), τ_e and τ_s denote effective correlation times for the internal motion with subscript *s* indicating that the correlation time is associated with the slower time scale, and R_{ex} denotes exchange contributions to R_2; in addition, the correlation time for overall rotational diffusion, τ_c, was also fitted.[48] Side-chain methyl-axis model-free optimization was performed using in-house routines implemented in Matlab. Three different models were fitted using two {O^2, τ_f}, three {O^2, τ_f, τ_{eff}}, or four {O_f^2, O_s^2, τ_f, τ_{eff}} parameters, where τ_f is associated with fast motions, $\tau_{eff} = (1/\tau_c + 1/\tau_s)^{-1}$ and τ_s denotes the correlation time for slow internal motions on par with τ_c.[49] The global correlation time was fixed to the value obtained from the backbone model-free optimization. Model selection was performed using an *F*-test at the level α = 0.95 ($p < 0.05$).[50]

Conformational entropy estimates from order parameters. The backbone conformational entropy change, going from state A to B, was estimated from the NMR order parameters using the relationship:[1,51]

$$\Delta S_{AB} = R \sum_k ln\left(\frac{1-O_{B,k}^2}{1-O_{A,k}^2}\right) \quad (1)$$

where $O_{X,k}$ is the order parameter for residue k in state X, and the sum runs over all residues. In a similar way, the conformational entropy change of the side chain methyl-axis was determined using:[51]

$$\Delta S_{AB} = R \sum_m C_m \sum_n \left(O_{B,n}^2 - O_{A,n}^2\right) \quad (2)$$

where C_m is a function of the residue type. The sums run over all residues n of type m. C_m = 1.32 for Val and Thr, 3.1 for Ile and Leu, and 2.31 for Met. The entropy for Ala side chains were calculated using eq (1).

Molecular dynamics simulations and analysis. All MD simulations were run with the Amber 14 software suite.[52] The X-ray crystal structures of the S–galectin-3C and R–galectin-3C complexes were used as the starting points for MD simulations. The PDB structure 3ZSL was used for the simulations of apo galectin-3C. Separate simulations were run for the two different conformations observed for ligand S. All crystal-water molecules were kept in the simulations. Each galectin-3C complex was solvated in an octahedral box of water molecules extending at least 10 Å from the protein using the tleap module, so that 4965–5593 water molecules were included in the simulations. The simulations were set up in the same way as in our previous studies of galectin-3C.[4,53,54] All Glu and Asp residues were assumed to be negatively charged and all Lys and Arg residues positively charged, whereas the other residues were neutral. The active-site residue His158 was protonated on the ND1 atom, whereas the other three His residues were protonated on the NE2 atom, in accordance with NMR measurements and previous extensive test calculations with MD.[55] This resulted in a net charge of +4 for the protein. No counter ions were used in the simulations.

The protein was described by the Amber ff14SB force field,[56] water molecules with the TIP4P-Ewald model,[57] whereas the ligands were treated with the general Amber force field.[58] Charges for the ligands were obtained with the restrained electrostatic potential method.[59] The ligands were optimized with the semiempirical AM1 method, followed by a single-point calculation at the Hartree–Fock/6-31G* level to obtain the electrostatic potentials, sampled with the Merz–Kollman scheme.[60] These calculations were performed with the Gaussian 09 software.[61] The potentials were then used by antechamber to calculate the charges. A few missing parameters were obtained with the Seminario approach:[62] The geometry of the ligands was optimized at TPSS/def2-SV(P) level, followed by a frequency calculation using the aoforce module of Turbomole 7.01.[63] From the resulting Hessian matrix, parameters for the missing angles and dihedrals were extracted with the Hess2FF program.[64] These parameters are given in Table S1 in the supplementary material.

For each complex, 10,000 steps of minimization were used, followed by 20 ps constant-volume equilibration, and 20 ps constant-pressure equilibration, all performed with heavy non-water atoms restrained towards the starting structure with a force constant of 209 kJ/mol/Å^2. Finally, the system was equilibrated for 2 ns, followed by 10 ns of production simulation, both performed with constant pressure and without any restraints (still with restraints for the GIST analysis). For each protein–ligand complex, 10 independent simulations were run, employing different solvation boxes and different starting velocities.[65] Consequently, the total simulation time for each complex was 100 ns. All bonds involving hydrogen atoms were constrained to the equilibrium value using the SHAKE algorithm,[66] allowing for a time step of 2 ps. The temperature was kept constant at 300 K using Langevin dynamics,[67] with a collision frequency of 2 ps^{-1}. The pressure was kept constant at 1 atm using a weak-coupling isotropic algorithm[68] with a relaxation time of 1 ps. Long-range electrostatics were handled by particle-mesh Ewald (PME) summation[69] with a fourth-order B spline interpolation and a tolerance of 10^{-5}. The cut-off radius for Lennard–Jones interactions between atoms of neighboring boxes was set to 8 Å. The snapshots were analyzed with the cpptraj module.[35]

Conformational entropy estimates from MD simulations. To validate the MD trajectories by NMR, we calculated order parameters from the MD trajectories. The N–H order parameters were obtained using isotropic reorientational eigenmode dynamic analysis.[70] The covariance matrix of the NH bond vectors was obtained from the trajectories by the cpptraj module[35] in the Amber 14 software.[52]

A total of 10,000 snapshots with a 10 ps sampling frequency were used for entropy and order parameter estimates, employing separate simulations for the complexes, for free galectin-3C and for the solvated ligands. Conformational entropies were calculated from the ensemble of configurations of the protein and ligands by analyzing the dihedral angle fluctuations.[54,71–73] The Cartesian coordinates from the trajectories were transformed to internal coordinates and the entropies were then calculated from probability distributions over all possible states of these coordinates using a bin size of 5° (i.e., 72 bins per dihedral). Entropies were normalized to that of a free rotor.[73] All entropies are reported as $-T\Delta S$ at 300 K.

Both entropies and order parameters were calculated as averages over 50 simulations of 2 ns each (with 200 snapshots in each, i.e. each of the 10 simulations were divided into five parts of equal length). The 2-ns time window is similar to the rotational correlation time of the protein. This procedure yields more stable entropy estimates by restricting the dependence on rare events.[74] The reported uncertainties are standard errors over these 50 simulations.

Water structure and solvation thermodynamics. We analyzed the structure and thermodynamics of the solvent around the two ligands (R and S) bound to galectin-3C, using GIST,[75] implemented in the *cpptraj* module of the Amber 14 software. The analysis was based on 80,000 snapshots from 10 independent 10-ns long trajectories for each complex in which the protein was kept restrained towards the starting crystal structure (the initial 2 ns of the simulations were not included in the analysis). The water–water interaction energy, E_{ww}, and solute–water interaction energy, E_{sw}, as well as translational, S_{trans}, and rotational, S_{rot}, entropy contributions were calculated for a rectangular grid of dimensions 54 Å × 27 Å × 30 Å,

centered on the ligand and extended at least 3 Å on each side of the ligand. The grid was divided into cubic boxes (0.5 Å × 0.5 Å × 0.5 Å), for which the thermodynamic properties were calculated. The sum of these properties over the entire region reveals the changes in the hydration thermodynamics of the region, relative to the thermodynamics of the bulk water.

RESULTS AND DISCUSSION

Ligand design and synthesis. We investigated the driving forces underlying affinity and selectivity in ligand binding by carrying out a comparative analysis involving two diastereomeric ligands R and S. The design of ligands R and S was inspired by the high-affinity (K_d = 2 nM) galectin-3 ligand 1-1'-sulfanediyl-bis-{3-deoxy-3-[4-(3-fluorophenyl)-1*H*-1,2,3-triazol-1-yl]-β-D-galactopyranoside}[20,76] (Fig. 1A). The high-affinity ligand interacts with galectin-3 via one of the galactose residues (that on the left-hand-side in Fig. 1A) in the conserved galactose binding site and the fluorophenyltriazolyl moieties interacts via face-to-face stacking with arginine side chains and one fluorine–amide orthogonal multipolar interaction.[20] The second galactose moiety interacts with only a single hydrogen bond (Fig. 1A; the non-interacting parts are depicted in gray) to the protein, leading us to hypothesize that this galactopyranose ring could be mimicked by a 2-hydroxypropyl chain, which would open up for the synthesis of two diastereomeric ligands R and S (Fig. 1B).

Synthesis of the ligands R and S relied on fine-tuning the reactivity between a 1-sulfhydryl-galactopyranose nucleophile and a doubly electrophilic glycidyl derivative: *In situ* fluoride-mediated activation of the masked nucleophilic triisopropylsilyl thio-galactoside and (R)- and (S)-glycidyl nosylate, respectively, proceeded stereoselectively in high yields, while other galactose nucleophiles (-SAc, -SH, thiouronium salts and thioxanthate) and glycidyl electrophiles (glycidyl tosylate, *tert*-butyl dimethyl silyl glycidyl and epi-chlorohydrin) gave lower yields and stereochemical scrambling due to nucleophilic attack occurring on both C1 and C3 of the glycidyl derivatives, or due to epoxide opening followed by intramolecular substitution to epoxide re-closing.

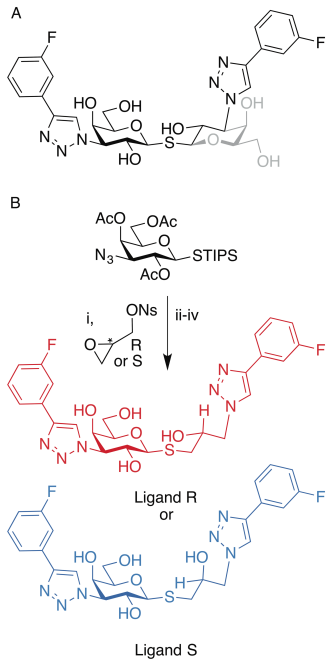


Figure 1. Chemical structures and synthesis of ligands. A, Chemical structure of the parent, high-affinity ligand 1-1'-sulfanediyl-bis-{3-deoxy-3-[4-(3-fluorophenyl)-1*H*-1,2,3-triazol-1-yl]-β-D-galactopyranoside}. The non-interacting atoms of one galactopyranose are depicted in gray. B, Synthesis and structures of the two ligands R (red) and S (blue). The stereo-center is located at the propyl C2. Reagents and conditions: i) TBAF*3H_2O, dry THF. ii) NaN_3, NH_4Cl, diglyme/H_2O 1:1. iii) 1-Ethynyl-3-fluorobenzene, CuI, Et_3N, DMF; iv) MeONa, MeOH.

Overall binding thermodynamics. We characterized the thermodynamics of ligand binding using ITC. We carried out five replicate titrations for each of ligands R and S, and analyzed the binding isotherms by performing a combined fit of the replicate data sets (Fig. 2; Fig. S1). Table 1 lists the resulting binding thermodynamics. Both ligands have dissociation constants in the low micromolar range, K_d(R) = (1.0 ± 0.03)×10^{-6} M and K_d(S) = (2.1 ± 0.1)×10^{-6} M, and the results correlate well with those obtained in competitive fluorescence polarization experiments, K_d(R) = (0.43 ± 0.04)×10^{-6} M and K_d(S) = (0.67 ± 0.5)×10^{-6} M. As reported previously,[5] K_d values determined by ITC are typically found to be higher by a factor of 2–4 than those measured by fluorescence polarization, but the relative affinities are unchanged within errors. The free energies of binding differ by only 1.9 ± 0.1 kJ/mol, but the differences in $\Delta H°_{tot}$ and $-T\Delta S°_{tot}$ are greater and consequently opposite in sign, indicating enthalpy–entropy compensation: $\Delta\Delta H°_{tot}$ = –5 ± 1 kJ/mol and $-T\Delta\Delta S°_{tot}$ = 3 ± 1 kJ/mol.

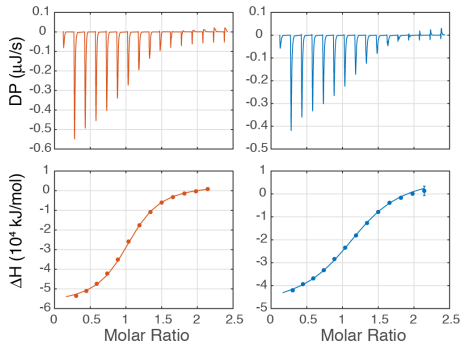


Figure 2. ITC experiments of ligand binding to galectin-3C. Example isotherms describing the titration of galectin-3C with (A) ligand R and (B) ligand S. The top panels show the raw thermograms of differential power plotted versus the ligand to protein molar ratio, while the lower panels show the resulting isotherms. The binding curve results from global fitting of 5 replicate data sets.

Crystal structures reveal subtle differences in binding modes. The crystal structures of the R– and S–galectin-3C complexes were refined to resolutions of 1.34 and 1.19 Å, respectively. As shown in Fig. 3, the two complexes have closely similar structures, with essentially no difference in the protein backbone conformation. The RMS deviation between the two structures is 0.13 Å for 473 backbone atoms and 0.59 Å when 2054 atoms are compared, including side chains.

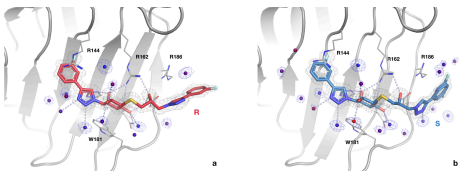


Figure 3. X-ray crystal structures of the ligand–galectin-3C complexes. A, R–galectin-3C. B, S–galectin-3C. The protein backbone is shown in ribbon representation (gray), key ligand-coordinating side chains are shown as lines, and hydrogen bonds are shown as dashed lines. The electron density of the ligand and water molecules is shown as gray mesh. Carbon atoms of the R ligand are colored red, while those of the S ligand are sky blue.

Below we will denote the aromatic ring substituents on galactose C3 as the "left hand side" (LHS), while the aromatic rings connected to the propylic chain will be referred as the "right hand side" (RHS); this notation is according to the viewpoint of Fig. 3 and all subsequent renditions of the structures. The LHS shows perfect overlap between the two complexes. The 3-fluorophenyl substituent sits in a pocket generated by the displacement of Arg144, with the fluorine atom pointing towards the protein backbone. There is no significant difference in the arrangement of water molecules around the LHS of the ligand in the two complexes.

Although R and S have a different configuration at propyl C2, the conformation of the ligand adjusts such that the hydroxyl group of the stereo-center maintains a highly similar hydrogen bond with Glu184. In R–galectin-3C, the C2 hydroxyl is positioned identically to the corresponding hydroxyl in the glucose moiety of natural ligands. In contrast, the adjustment of the C2 hydroxyl in the S ligand to maintain the hydrogen bond positions the propyl chain differently from the ring carbons of glucose, resulting in different interactions of the two ligands with the protein at the RHS of the binding site. Furthermore, the RHS of R is modeled with a single conformation, whereas the RHS of S is modeled as two conformations in which the fluorinated ring has two orientations related by an 180° flip. However, the conformational mobility of R in the crystal might be restricted compared to solution as the neighboring symmetry-related molecule in the crystal is very close. At the RHS, both R and S interact with Arg186, despite the differences in conformation at this end of the ligand. S appears to have tighter interaction with Arg186 due to a better alignment between the π orbitals of the ligand phenyl ring and the face of the arginine guanidine group.

For both complexes, the *B*-factors of the ligand atoms on the LHS are lower compared to those of the RHS, indicating that the LHS is more ordered. Furthermore, the *B*-factors of the LHS are almost identical in the two complexes. The electron density for Arg144, which stacks with the fluorinated phenyl ring of the LHS, is slightly less well-ordered in R than in S, but given that the ligands are chemically identical in this region and the water structure almost identical, this is unlikely to be significant. This result suggests that the LHS region does not contribute to the difference in free energy of binding the two diastereomers.

Water molecules are well conserved around the binding site. In particular, we see that waters around the LHS overlap almost perfectly between the two complexes, while for the RHS the different conformation of R allows an additional water molecule to sit between the ligand and the protein.

Chemical shift mapping of ligand binding. Chemical shift assignments of R–galectin-3C and S–galectin-3C were based on a HNCACB experiment and the apo galectin-3C assignments reported previously.[4,5] Minor chemical shift differences are observed for the backbone amides throughout the protein; the RMS chemical shift difference between the ligand-bound and apo forms of galectin-3C in the ^{1}H and ^{15}N dimensions are 0.05 ppm and 0.26 ppm for S–galectin-3C, and 0.06 ppm and 0.30 ppm for R–galectin-3C (Fig. S3a). The methyl chemical shifts show changes similar to those of the backbone, with RMSDs of 0.03 ppm (^{1}H) and 0.1 ppm (^{13}C) for S–galectin-3C, and 0.04 ppm and 0.1 ppm for R–galectin-3C (Fig. S3b). The largest chemical shift changes induced by ligand binding are observed for residues in close proximity to the ligand in the crystal structure (Fig. S3c), demonstrating that the binding mode observed in the crystal structure is maintained in solution.

Significant chemical shift differences between the R– and S–galectin-3C complexes are observed in key regions of the structure (Fig. 4). The overall chemical shift RMSD is 0.02 ppm and 0.14 ppm for backbone ^{1}H and ^{15}N, respectively, and 0.04 ppm and 0.02 ppm for methyl ^{1}H and ^{13}C, respectively. Two methyl groups show a weighted chemical shift difference greater than

0.05 ppm between the two complexes: Val170γ1 and Val172γ1. Furthermore, the ^{1}H and ^{15}N chemical shifts of the Arg162 and Arg186 guanidine groups differ between the two complexes. These side chains are located closely together and in proximity of the stereo-center of the ligand. Notably, chemical shift differences are observed also in regions of the protein where the average structures are virtually identical between the two complexes. This observation indicates that subtle differences exist in the conformational ensembles sampled by the two complexes, which is also confirmed by ensemble refinement of crystal structures, described next.

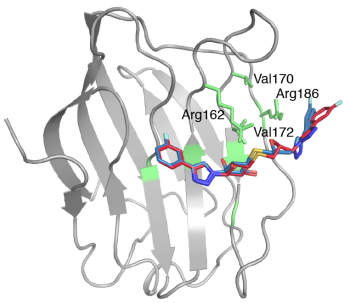


Figure 4. Chemical shift differences between the R– and S–galectin-3C complexes. Residues with weighted chemical shift differences | $\Delta\delta(R–S)| \geq 0.05$ ppm are highlighted in green on the structure of the R–galectin-3C complex with ligand S superimposed. These include methyl groups of Val170 and Val172 and guanidine groups of Arg162 and Arg186 in binding site. Val172 is situated beneath the side chain of Arg162 in the view of the figure. Weighted chemical shift differences were evaluated as $([\Delta\delta(^{1}H)]^2 + 0.1[\Delta\delta(^{15}N)]^2)^{1/2}$ for amides and $([\Delta\delta(^{1}H)]^2 + 0.25[\Delta\delta(^{13}C)]^2)^{1/2}$ for methyls.

Ensemble refinement of crystal structures highlight differences in flexibility. To investigate the conformational mobility of each complex in the crystal we carried out ensemble refinement of the structure against the X-ray diffraction data. The resulting ensembles indicate that the S diastereomer shows larger fluctuations in the crystal than does R, due to a large variation in the RHS sp^3 dihedral angles, as shown in Fig. 5. This result agrees with the dual conformation of ligand S observed in the traditionally refined crystal structure, although the conformational variation is much greater in the ensemble view. The ensemble refinement also confirms that the R ligand stays in a single conformation, although with some translational movement of the RHS end.

The ensemble-refined crystal structure of S–galectin-3C also shows a higher fluctuation of the protein backbone, especially in the Asn164–Arg168 loop region (Fig. 5), neighboring the RHS of the bound ligand. The side chains of Asn160 and Glu165, which form hydrogen bonds with both ligands, have well-defined positions, whereas larger fluctuations are observed for Arg144 and Arg168 that interact through π–π stacking with the two phenyl rings of ligands R and S. In fact, Arg168 mainly samples conformations in which it does not interact with the ligand at all, particularly so in S–galectin-3C. The great variability in the side-chain orientation of Arg144 is also reflected by the NMR data, see below.

The resulting ensembles indicate that the S–galectin-3C complex shows considerably higher variability than does R–galectin-3C, providing qualitative evidence that protein and ligand conformational entropy is greater in S-galectin-3C. Our attempts to quantitate the entropy difference from the ensembles resulted in calculated values that were consistent with our other results, but the estimated errors were far greater than the difference between the R- and S-complexes (data not shown).

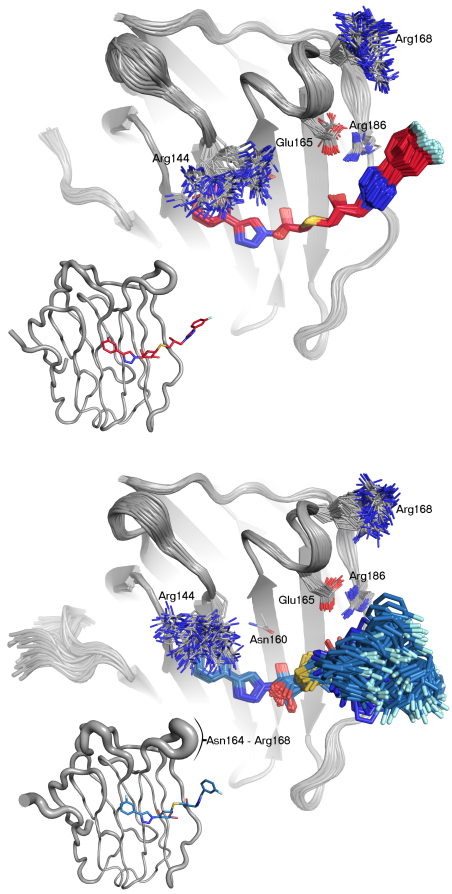


Figure 5. Ensemble refined X-ray crystal structures. Overlay of the best 100 structures generated by ensemble refinement for (A) R–galectin-3C (red) and (B) S–galectin-3C (blue). Insets: The protein backbone is displayed as a tube with a diameter corresponding to the RMS atom coordinate fluctuations in the ensemble.

Differences in conformational fluctuations measured by NMR. We carried out a suite of NMR relaxation experiments that probe conformational dynamics on the picosecond to nanosecond time scale to yield the amplitudes of conformational fluctuations in terms of order parameters, denoted O^2. We measured ^{15}N backbone relaxation rates at three static magnetic field strengths and methyl ^{2}H relaxation rates at two static magnetic field strengths. Out of 138 residues, ^{15}N relaxation data could be measured for 101 and 100 backbone amides in R–galectin-3C and S–galectin-3C, respectively. Likewise, out of a total of 85 methyl groups, ^{2}H relaxation rates could be measured for 65 and 47 methyl groups in R–galectin-3C and S–galectin-3C, respectively. The missing residues had cross-peaks that were overlapped or too broadened to allow for quantitative analysis.

We characterized the amplitudes of conformational fluctuations using the model-free formalism.[77,78] The best-fit rotational diffusion tensor is anisotropic with a correlation time (τ_c) of 7.5 ns and 8.1 ns, anisotropy of 1.1 and 1.1, and rhombicity of 0.9 and 1.2 for R–galectin-3C and S–galectin-3C, respectively. The higher value of τ_c observed for S–galectin-3C is fully explained by the slightly higher concentration of DMSO in this sample, which increases the solvent viscosity.[79,80]

The backbone order parameters are very similar in the two complexes; the mean values are $\langle O^2 \rangle = 0.85 \pm 0.05$ and 0.84 ± 0.05 for R–galectin-3C and S–galectin-3C, respectively. A significant difference in O^2 is observed for residues Tyr118, Ile132, Ile171, Asp178, Arg183, and Leu242, all of which are rather peripheral with respect to the binding site (Fig. 6). This result indicates that the different stereochemistry of the ligand and the associated differences in protein conformation affect the amplitudes of backbone fluctuations at remote locations.

The order parameters for the methyl-bearing side chains vary significantly over the protein. However, the differences between the two complexes are overall small, except for residues Val127, Ile132, Val172, Val189, Ala216, Leu228 and Ile231, which show larger deviations. Out of these residues, only Val172 is located in the binding site, right next to the stereo-center. The side chain of Val172 shows a larger degree of freedom in the R–galectin-3C complex.

Arginine side chains play a special role in ligand coordination by galectin-3C. Arg144, Arg162, and Arg186 form close interactions with the ligand (cf. Fig. 3). However, Arg144 is not observed in the NMR spectra, presumably as a consequence of intermediate exchange of the side-chain guanidine group between alternative positions. This result is in agreement with the ensemble-refined crystal structures, in which Arg144 shows extensive flexibility. ^{15}N side-chain order parameters could be measured for 5 out of 9 arginines. Only Arg129 and Arg224 show a significant difference between the two complexes, and both of these are located peripherally to the binding site.

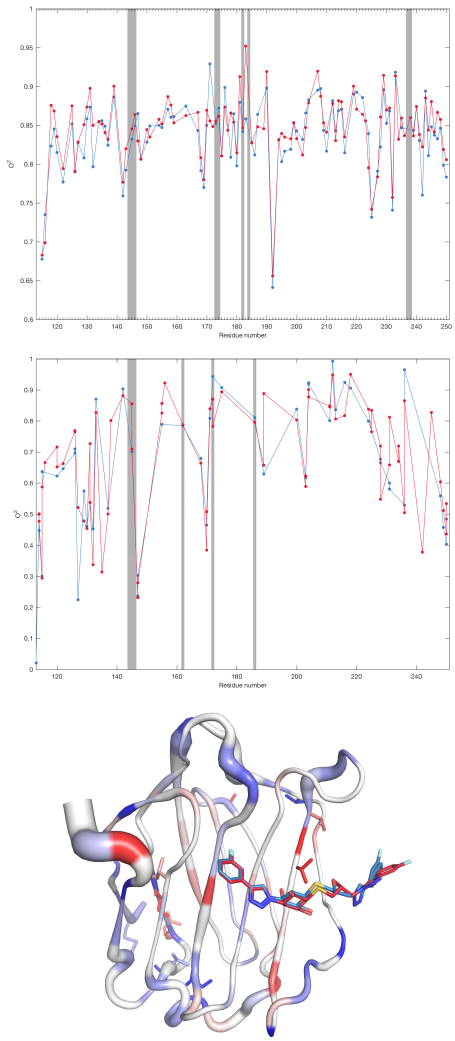


Figure 6. NMR order parameters for R– and S–galectin-3C. A, Backbone O^2 values. B, Side chain O^2 values for arginine ^{15}Nε and methyl axes. Data for R– and S–galectin-3C are shown in red and blue, respectively. Gray bars indicate residues in contact with the ligand (residues for which any backbone amide atom or methyl atom is within 5 Å of any ligand atom). C, ΔO^2 color coded onto the R–galectin-3C structure with ligand S superimposed. Residues with ΔO^2(R–S) > 0 are colored blue, while those with ΔO^2(R–S) < 0 are colored red. The intensity of the color scales with the magnitude of ΔO^2 from red via pink to white ($-0.1 \leq \Delta O^2$(R–S) < 0) and from white via light

blue to dark blue ($0 < \Delta O^2(\text{R–S}) \leq 0.1$). Side chains are shown in stick representation for residues with a difference in side-chain order parameters of $|\Delta O^2(\text{R–S})| > 0.05$. The width of the tube indicates the average backbone O^2 values in the two complexes: a wider tube indicates a lower order parameter and vice versa.

Differences in conformational fluctuations determined by MD simulations. To sample conformational fluctuations on a time scale shorter than, or on par with overall rotational diffusion, we performed MD simulations on both complexes. Since the crystal structures of the ligand–galectin-3C complexes show two conformations of S (but only one of R), we initiated separate MD simulations for the two conformers. We studied how the conformation of the ligand varied in the MD simulations by following the dihedral angle representing the orientation of the RHS phenyl ring (ϕ), shown in Fig. S4a). In each of the three trajectories, the ligand samples a unimodal and equally wide (~50º) distribution of the ϕ dihedral, indicating that the rotation barrier is high enough that the ligand does not change conformation on the nanosecond time scale (Fig. S4b).

Conformational entropy differences estimated by NMR. Based on the experimental order parameters, we estimated the difference in the conformational entropy between the two complexes, see eqs. [1] and [2]. Despite the average values of O^2 being highly similar, the residue-specific differences lead to a significant difference in backbone conformational entropy between galectin-3C in the R- and S-bound states, $-T\Delta\Delta S_{bb}$(R–S) = 17 ± 5 kJ/mol. By contrast, the corresponding result for the methyl-axis O^2 is not statistically significant: $-T\Delta\Delta S_{sc}$(R–S) = –5 ± 6 kJ/mol. Taken together, the NMR order parameters yield an estimate of $-T\Delta\Delta S_{bb+sc}$(R–S) = 12 ± 8 kJ/mol, indicating that galectin-3C in the R-bound state has lower conformational entropy than it has in the S-bound state (Table 2). That is, the conformational entropy difference between the two complexes has the same sign as, but a greater magnitude than $-T\Delta\Delta S_{tot}$(R–S) obtained by ITC, suggesting that conformational entropy makes a significant contribution to the overall binding thermodynamics. It should be noted that the NMR-based estimate covers only a subset of the dihedral angles in the protein. However, it serves as a useful reference for validating the MD simulations, which provide the total conformational entropy of both galectin-3C and the bound ligand.

Conformational entropy differences determined by MD. We calculated the conformational entropy of galectin-3C and the bound ligands in both complexes, using dihedral angle distributions from the MD simulations. Table 2 shows the difference in conformational entropy between the two ligand–galectin-3C complexes. For both complexes, the dihedral flexibility of galectin-3C decreases upon ligand binding (Table S1). The effect yields a change in conformational entropy, $-T\Delta S_{conf}$, of 43 kJ/mol and 32–33 kJ/mol for the protein in the R- and S-bound states, respectively (we run separate MD simulations of the two conformations of S observed in the crystal structure and they gave entropies that agreed within 1 kJ/mol). The decrease is largest for Arg186 for both ligands (7–8 kJ/mol), as can be seen in Fig. S5. This residue forms hydrogen bonds with Glu184, which interacts with the ligands R and S and shows the second largest decrease in entropy when ligand S binds (4 kJ/mol), but a rather small decrease upon binding ligand R (1 kJ/mol). Arg144 also gives a rather large negative entropy contribution upon binding either ligand (3–4 kJ/mol). Ile171 and Leu219 give large negative contributions (3–4 kJ/mol) when binding S, but smaller when binding R (1 kJ/mol). Significantly increased conformational entropy is observed for 9–12% of the residues upon ligand binding, with the largest contribution from His158 (2–3 kJ/mol) for both ligands.

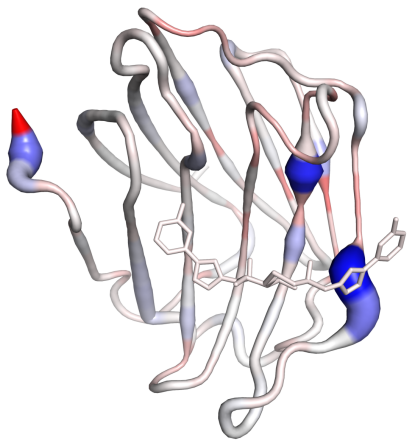


Figure 7. Conformational entropy contributions to $T\Delta\Delta S_{conf}$ (R–S), reported per residue. $T\Delta\Delta S_{conf}$ (R–S) is color coded onto the galectin-3C structure with blue hues indicating $T\Delta\Delta S_{conf}(\text{R–S}) > 0$ and red hues indicating $T\Delta\Delta S_{conf}(\text{R–S}) < 0$, with the color intensity ranging from weak (white) for $T\Delta\Delta S_{conf} = 0$ to intense (maximally blue or red) for $|T\Delta\Delta S_{conf}| = 3$ kJ/mol. The figure is based on the crystal structure of S–galectin-3C.

The total conformational entropy of the protein is greater for S–galectin-3C than for the R-complex, $-T\Delta\Delta S_{conf}$(R–S) = 11 ± 5 kJ/mol (taking into account both MD trajectories for S–galectin-3C), which is statistically significant at the 95% level. This result is in general agreement with the estimate obtained from NMR ($-T\Delta\Delta S_{conf}$(R–S) = 12 ± 8 kJ/mol), which includes only a subset of the dihedral angles for the protein. The difference between complexes arises from small contributions from many residues (Fig. 7). At the level of individual residues, 22–23% show a statistically significant contribution with the same sign as the total difference, whereas 9–11% show the opposite behavior. Among the latter, the largest contributions (–3 kJ/mol) come from Ile171 and Glu184, mentioned above (Fig. 7).

The change in conformational entropy of the ligand upon complex formation is 24 ±1 kJ/mol and 25–26 ± 1 kJ/mol for R– and S–galectin-3C, respectively. The difference between R and S is not statistically significant, neither in the bound nor in the free states. The indistinguishable conformational entropy of the free ligands is in line with the expectation that they should

have nearly identical chemical potential in the free state, based on their diastereomeric relationship.

Taken together, the difference in conformational entropy between the two complexes, $-T\Delta\Delta S_{conf}$(R–S) = 9 ± 5 kJ/mol, is greater than the difference in the net binding entropy, $-T\Delta\Delta S°_{tot}$(R–S) = 3 ± 1 kJ/mol, indicating that the difference in conformational entropy might be compensated by solvation entropy. Note that due to the design of the system, there should not exist any other degrees of freedom contributing to the entropy of binding.

Grid inhomogeneous solvation theory reveals key differences in solvation between the two complexes. The MD simulations of the R–galectin-3C and S–galectin-3C complexes reveal different hydration thermodynamics relative to bulk water (Table 3). The difference in total solvation free energy between the two complexes is $\Delta\Delta G_{solv}$(R–S) = –71 ± 3 kJ/mol, i.e., solvation is more favorable for the R–galectin-3C complex. $\Delta\Delta G_{solv}$(R–S) is completely dominated by the difference in solvation enthalpy, $\Delta\Delta H_{solv}$(R–S) = –74 ± 3 kJ/mol. More precisely, the solvation enthalpy is dominated by solute–water interactions, –79 ± 2 kJ/mol, whereas differences in water–water interactions are minor by comparison, 4 ± 2 kJ/mol. Within the binding site, the solute–water interaction energies differ primarily in the area where the two ligands bind differently, i.e., at the RHS (Fig. 8). A detailed analysis shows that protein–water interactions dominate the solute–water term, and are more favorable for the R–galectin-3C complex, –106 ± 2 kJ/mol, whereas contributions from the ligand–water interactions are four times smaller and favor the S–galectin-3C complex, +27 ± 1 kJ/mol. The difference between the two complexes in solvation entropy amounts to only $-T\Delta\Delta S_{solv}$(R–S) = 3 ± 1 kJ/mol, which is but a fraction (4%) of $\Delta\Delta H_{solv}$(R–S). The entropic contribution from solvation enhances the conformational entropy of the protein and ligand ($-T\Delta\Delta S_{conf}$(R–S) = 9 ± 5 kJ/mol). However, owing to the rather large uncertainties in both the experimental and calculated entropies, the net result ($-T\Delta\Delta S_{conf+solv}$(R–S) = 13 ± 5 kJ/mol) is still in agreement with the overall entropy determined by ITC, $-T\Delta\Delta S_{tot}$(R–S) = 3 ± 1 kJ/mol. Moreover, the large difference in solvation enthalpy is compensated the protein–protein and protein–solvent (outside the box) enthalpies, which are large and hard to estimate accurately, whereas the difference in the protein–ligand interaction energies between the R and S ligands is modest, –4.6 ± 0.2 kJ/mol.

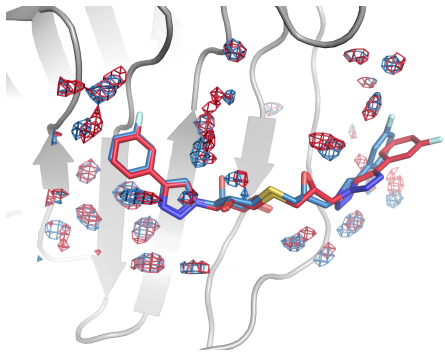


Figure 8. Differences in solvation around the binding site. Regions with favorable solute–water interactions relative to bulk water are represented as red (R–galectin-3C) and blue (S–galectin-3C) meshes. The cut-off threshold was set to –8.4 kJ/mol.

The GIST calculations confirm the presence of the extra water molecule (W85) situated between the ligand and protein in the crystal structure of R–galectin-3C, but not in S–galectin-3C (cf. Fig. 3).

Thus, we conclude that the higher binding affinity for the R diastereomer includes an important contribution from favorable hydration thermodynamics that is dominated by protein–water interactions around the binding site. Galectin-3 has a relatively solvent accessible binding site, which engages numerous water molecules, a feature that certainly contributes greatly to the present results. It would be of great interest to carry out future research to investigate other proteins with different types of binding sites, e.g., those that are less solvent accessible.[12]

Concluding remarks. We have carried out a comparative analysis of ligand binding to galectin-3C using two diastereomeric ligands and a range of experimental techniques combined with computational methods. This approach has the important advantage that any differences in the thermodynamics of the two binding processes can be largely related to the bound state, while the contributions from the free states are expected to cancel — as borne out by the present results. The two ligands exhibit closely similar free energies of binding, as might be expected for diastereomers and which potentially could limit their usefulness for the purpose of unraveling the statistical thermodynamics of protein–ligand binding. However, this system exhibits enthalpy–entropy compensation, so that the two complexes still manifest meaningful differences in both binding enthalpy and entropy that we investigated to pinpoint the driving forces underlying the thermodynamic signatures of binding. Thus, based on this experimental design, we were able to dissect the enthalpic and entropic contributions to ligand selectivity and affinity.

Our results show that the S diastereomer has a more flexible bonding to galectin-3C than the R isomer, which translates to a more favorable entropy, for the protein, ligand, as well as the

surrounding solvent. This is compensated by a more favorable enthalpy for the R isomer.

In a broader perspective, the sensitive interdependence of the enthalpy and entropy (from both solvent and protein) adds to our understanding of molecular recognition. The phenomenon indicates both opportunities and challenges in rational drug design. On the one hand, contributions from solvation entropy to the free energy of binding are well known, and the present results reiterate the concept to target individual water sites to achieve increased binding affinity.[12,81] On the other hand, efforts to design ligands that perturb the solvent structure around the binding site might not achieve the expected result due to changes in conformational entropy of the ligand and protein, as exemplified herein.

Author Contributions

MLV synthesized, purified, and characterized ligands. OS performed NMR and ITC experiments and analysis. FM crystallized complexes and determined their structures by X-ray diffraction together with MLV and DL. MO carried out MD simulations and entropy calculations. OC carried out ensemble refinement together with EO. MMI performed GIST calculations. HL, UJN, DL, EO, UR and MA designed the project, planned the experiments, and analyzed data. The manuscript was written through contributions of all authors. ‡These authors contributed equally.

ACKNOWLEDGMENT

We thank Ulrich Weininger and Göran Carlström for assistance with NMR experiments, Natalia Markova (Malvern) for helpful discussion on ITC measurements and analysis, and Malvern Instruments for access to ITC analysis software. Protein production was carried out by the Lund Protein Production Platform (LP3) at Lund University. High-field (900 MHz) NMR spectroscopy was carried out at the Swedish NMR Center at University of Gothenburg, supported by the Knut and Alice Wallenberg Foundation (NMR for Life). This work was supported by the Knut and Alice Wallenberg Foundation (KAW 2013.022), the Swedish Research Council (project 2014-5540 to UR) and the European Spallation Source (ERIC; UR and DL).

REFERENCES

(1) Akke, M.; Brüschweiler, R.; Palmer, A. G. *J. Am. Chem. Soc.* **1993**, *115*, 9832.
(2) Marlow, M. S.; Dogan, J.; Frederick, K. K.; Valentine, K. G.; Wand, A. J. *Nat. Chem. Biol.* **2010**, *6* (5), 352.
(3) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448* (7151), 325.
(4) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. *J. Biomol. NMR* **2009**, *45* (1–2), 157.
(5) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U. J.; Akke, M. *J. Am. Chem. Soc.* **2010**, *132*, 14577.
(6) Tzeng, S. R.; Kalodimos, C. G. *Nature* **2012**, *488* (7410), 236.
(7) Gill, M. L.; Byrd, R. A.; Palmer Arthur G., I. I. I. *Phys. Chem. Chem. Phys.* **2016**, *18* (8), 5839.
(8) MacRaild, C. A.; Daranas, A. H.; Bronowska, A.; Homans, S. W. *J. Mol. Biol.* **2007**, *368* (3), 822.
(9) Syme, N. R.; Dennis, C.; Bronowska, A.; Paesen, G. C.; Homans, S. W. *J. Am. Chem. Soc.* **2010**, *132* (25), 8682.
(10) Haider, K.; Wickstrom, L.; Ramsey, S.; Gilson, M. K.; Kurtzman, T. *J. Phys. Chem. B* **2016**, *120* (34), 8743.
(11) Fenley, A. T.; Muddana, H. S.; Gilson, M. K. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (49), 20006.
(12) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 808.
(13) Saraboji, K.; Håkansson, M.; Genheden, S.; Diehl, C.; Qvist, J.; Weininger, U.; Nilsson, U. J.; Leffler, H.; Ryde, U.; Akke, M.; Logan, D. T. *Biochemistry* **2012**, *51*, 296.
(14) Leffler, H.; Carlsson, S.; Hedlund, M.; Qian, Y. N.; Poirier, F. *Glycoconj. J.* **2002**, *19* (7–9), 433.
(15) Dumic, J.; Dabelic, S.; Flögel, M. *Biochimica et Biophysica Acta - General Subjects*. 2006, pp 616–635.
(16) Liu, F. T.; Rabinovich, G. A. *Year Immunol. 2* **2010**, *1183*, 158.
(17) Liu, F.-T.; Rabinovich, G. A. *Nat. Rev. Cancer* **2005**, *5* (1), 29.
(18) Di Lella, S.; Sundblad, V.; Cerliani, J. P.; Guardia, C. M.; Estrin, D. A.; Vasta, G. R.; Rabinovich, G. A. *Biochemistry* **2011**, *50* (37), 7842.
(19) Blanchard, H.; Yu, X.; Collins, P. M.; Bum-Erdene, K. *Expert Opin. Ther. Pat.* **2014**, *24* (10).
(20) Delaine, T.; Collins, P.; MacKinnon, A.; Sharma, G.; Stegmayr, J.; Rajput, V. K.; Mandal, S.; Cumpstey, I.; Larumbe, A.; Salameh, B. A.; Kahl-Knutsson, B.; van Hattum, H.; van Scherpenzeel, M.; Pieters, R. J.; Sethi, T.; Schambye, H.; Oredsson, S.; Leffler, H.; Blanchard, H.; Nilsson, U. J. *ChemBioChem* **2016**, *17* (18), 1759.
(21) Mandal, S.; Nilsson, U. J. *Org. Biomol. Chem.* **2014**, *12* (27), 4816.
(22) Keller, S.; Vargas, C.; Zhao, H.; Piszczek, G.; Brautigam, C. A.; Schuck, P. *Anal. Chem.* **2012**, *84* (11), 5066.
(23) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes. The Art of Scientific Computing*; Cambridge University Press: Cambridge, 1986.
(24) Freiburger, L.; Auclair, K.; Mittermaier, A. *Methods* **2015**, *76*, 149.
(25) Salomonsson, E.; Larumbe, A.; Tejler, J.; Tullberg, E.; Rydberg, H.; Sundin, A.; Khabut, A.; Frejd, T.; Lobsanov, Y. D.; Rini, J. M.; Nilsson, U. J.; Leffler, H. *Biochemistry* **2010**, *49* (44), 9518.
(26) Ursby, T.; Unge, J.; Appio, R.; Logan, D. T.; Fredslund, F.; Svensson, C.; Larsson, K.; Labrador, A.; Thunnissen, M. M. *J Synchr Radiat.* **2013**, *20* (Pt 4), 648.
(27) Kabsch, W. In *International Tables for Crystallography, Volume F: Crystallography of Biological Macromolecules*; Rossmann, M. G., Arnold, E., Eds.; Kluwer Academic Publishers: Dordrecht, 2010; Vol. 66, pp 125–132.
(28) Evans, P. R.; Murshudov, G. N. *Acta Crystallogr D Biol Crystallogr* **2013**, *69* (Pt 7), 1204.
(29) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 4), 355.
(30) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J Comput Chem* **2004**, *25* (13), 1605.
(31) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr D Biol Crystallogr* **2009**, *65* (Pt 10), 1074.
(32) Afonine, P. V; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr D Biol Crystallogr* **2012**, *68* (Pt 4), 352.

(33) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 4), 486.
(34) Burnley, B. T.; Afonine, P. V.; Adams, P. D.; Gros, P. *Elife* **2012**, *1*, e00311.
(35) Roe, D. R.; Cheatham III, T. E. *J Chem Theory Com* **2013**, *9* (7), 3084.
(36) Wittekind, M.; Mueller, L. *J. Magn. Reson. Ser. B* **1993**, *101*, 201.
(37) Bax, A.; Clore, G. M.; Driscoll, P. C.; Gronenborn, A. M.; Ikura, M.; Kay, L. E. *J. Magn. Reson.* **1990**, *87*, 620.
(38) Bax, A.; Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1990**, *88*, 425.
(39) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(40) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, P.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins-Structure Funct. Bioinforma.* **2005**, *59*, 687.
(41) Loria, J. P.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 2331.
(42) Skrynnikov, N. R.; Mulder, F. A. A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 4556.
(43) Millet, O.; Muhandiram, D. R.; Skrynnikov, N. R.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 6439.
(44) Ahlner, A.; Carlsson, M.; Jonsson, B. H.; Lundström, P. *J. Biomol. NMR* **2013**, *56* (3), 191.
(45) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2003**, *25*, 25.
(46) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2008**, *40* (2), 121.
(47) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2008**, *40* (2), 107.
(48) Mandel, A. M.; Akke, M.; Palmer, A. G. *J. Mol. Biol.* **1995**, *246*, 144.
(49) Skrynnikov, N. R.; Millet, O.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 6449.
(50) Devore, J. L. *Probability and Statistics for Engineering and the Sciences*, 5th ed.; Brooks/Cole Publishing Company: Monterey, 1999.
(51) Li, D. W.; Brüschweiler, R. *J. Am. Chem. Soc.* **2009**, *131* (21), 7226.
(52) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; others. *University of California, San Francisco*. 2015,.
(53) Genheden, S.; Diehl, C.; Akke, M.; Ryde, U. *J. Chem. Theory Comput.* **2010**, *6* (7), 2176.
(54) Genheden, S.; Akke, M.; Ryde, U. *J. Chem. Theory Comput.* **2014**, *10* (1), 432.
(55) Uranga, J.; Mikulskis, P.; Genheden, S.; Ryde, U. *Comput. Theor. Chem.* **2012**, *1000*, 75.
(56) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11*, 3696.
(57) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120* (20), 9665.
(58) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem* **2004**, *25*, 1157.
(59) Bayly, C. C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. a. *J. Phys. Chem.* **1993**, *97* (40), 10269.
(60) Besler, B.; Merz, K.; Kollman, P. *J. Comput.* **1990**.
(61) Gaussian09, R. A. *Inc., Wallingford CT* **2009**.
(62) Seminario, J. M. *Int. J. Quantum Chem.* **1996**, *60*, 1271.
(63) Turbomole, V. 0. *There is no Corresp. Rec. this Ref.* **2015**.
(64) Nilsson, K.; Lecerof, D.; Sigfridsson, E. *Sect. D Biol. …* **2003**.
(65) Genheden, S.; Ryde, U. *J. Comput. Chem.* **2011**.
(66) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23* (3), 327.
(67) Wu, X.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381* (3–4), 512.
(68) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. van; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(69) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98* (12), 10089.
(70) Prompers, J. J.; Brüschweiler, R. *J. Am. Chem. Soc.* **2002**, *124*, 4522.
(71) Edholm, O.; Berendsen, H. J. C. *Mol. Phys.* **1984**, *51* (4), 1011.
(72) Trbovic, N.; Cho, J. H.; Abel, R.; Friesner, R. A.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2009**, *131* (2), 615.
(73) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. *J. Biomol. NMR* **2009**, *45* (1–2), 157.
(74) Genheden, S.; Akke, M.; Ryde, U. *J. Chem. Theory Comput.* **2013**, *in press*.
(75) Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. *J. Chem. Phys.* **2012**, *137* (4), 973.
(76) Peterson, K.; Kumar, R.; Stenstrom, O.; Verma, P.; Verma, P. R.; Hakansson, M.; Kahl-Knutsson, B.; Zetterberg, F.; Leffler, H.; Akke, M.; Logan, D. T.; Nilsson, U. J. *J. Med. Chem.* **2018**.
(77) Halle, B.; Wennerström, H. *J. Chem. Phys.* **1981**, *75*, 1928.
(78) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(79) Schichman, S. A.; Amey, R. L. *J. Phys. Chem.* **1971**, *75* (6), 98.
(80) Catalan, J.; Diaz, C.; Garcia-Blanco, F. *J. Org. Chem.* **2001**, *66* (17), 5846.
(81) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2817.

Table 1. Overall binding thermodynamics from ITC

Complex	K_d (10^{-6} M)	$\Delta G°_{tot}$ (kJ/mol)	$\Delta H°_{tot}$ (kJ/mol)	$-T\Delta S°_{tot}$ (kJ/mol)
R–galectin-3C	1.0 ± 0.03	–34.6 ± 0.1	–60.4 ± 0.4	25.8 ± 0.4
S–galectin-3C	2.1 ± 0.1	–32.7 ± 0.1	–55.7 ± 0.9	22.9 ± 0.9
Difference R–S	NA	–1.9 ± 0.1	–5 ± 1	3 ± 1

a Table footnote here.

Table 2. Conformational entropy differences between R– and S–galectin-3C

Method	$-T\Delta\Delta S$ (kJ/mol)
NMR[a]	12 ± 8
MD[b]	11 ± 5
MD[c]	9 ± 5

[a] Includes protein dihedrals of the backbone and methyl-bearing side chains.

[b] Includes all protein dihedrals.

[c] Includes all protein and ligand dihedrals.

Table 3. Solvation thermodynamics from GIST calculations.[a]

Complex	R–galectin-3C	S–galectin-3C	Difference R–S
$-T\Delta S_{rot}$	225.2±0.6	218.3±0.5	6.9±0.8
$-T\Delta S_{trans}$	183.3±0.4	187.1±0.3	–3.7±0.5
$-T\Delta S_{solv}$	408.6±0.8	405±0.6	3.2±1.0
ΔE_{p-w}	–1463.2±1.3	–1357.5±1.4	–105.7±2.0
ΔE_{l-w}	–259.2±0.4	–286.3±0.4	27 ± 1
ΔE_{s-w}	–1722.4±1.4	–1643.8±1.5	–78.6±2.0
ΔE_{w-w}	–3221.5±1.9	–3225.7±1.2	4.1±2.2
ΔE_{solv}	–4943.9±1.8	–4869.5±2.0	–74.5±2.7
ΔG_{solv}	–4535.3±1.5	–4464.0±2.2	–71.3±2.6

[a] Rotational, ΔS_{rot}, and translational, ΔS_{trans}, entropy as well as the protein–water interaction energy, ΔE_{p-w}, ligand–water interaction energy, ΔE_{l-w}, and water–water interaction energy, ΔE_{w-w}, of the studied region, shown relative to bulk water. $\Delta S_{solv} = \Delta S_{rot} + \Delta S_{trans}$, $\Delta E_{s-w} = \Delta E_{p-w} + \Delta E_{l-w}$, $\Delta E_{solv} = \Delta E_{s-w} + \Delta E_{w-w}$, and $\Delta G_{solv} = \Delta E_{solv} - T\Delta S_{solv}$. All terms are in kJ/mol. Reported uncertainties are the standard errors over the ten independent MD simulations.

Table S1. Conformational entropy differences between the various R– and S–galectin-3C complexes and the apo protein, obtained from the MD simulations[a]

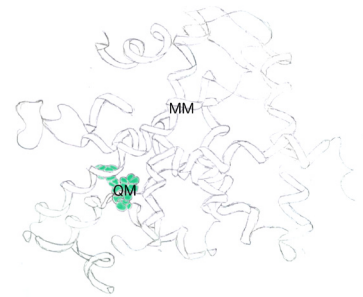
	R–apo	S–apo	S2–apo
$-T\Delta\Delta S$ [b]	43 ± 5	33 ± 5	32 ±5
$-T\Delta\Delta S$ [c]	67 ± 5	57 ± 4	58 ± 5

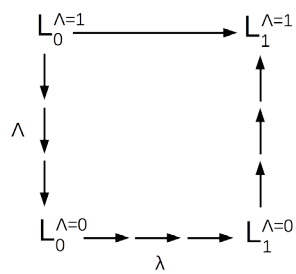
[a] S2 is the second conformation of the S ligand in complex with galectin-3C.

[b] Includes all protein dihedrals.

[c] Includes all protein and ligand dihedrals.

Proteins are one of the building blocks of the human body. They have a variety of important and necessary biological functions, participating in chemical reactions and processes. Most proteins have one or several binding pockets and binding of other molecules controls the biological function of the protein. This function may in some cases be linked to a disease. With organic molecules that effectively bind to the binding pocket of a protein, the biological function of the protein may be modified. In such cases, the organic molecule may become a drug. A critical factor determining whether an organic molecule is better than another organic molecule to bind to a protein is the difference in binding affinities between the protein and the organic molecules. So far, most drug discovery has been conducted through experimental studies, involving synthesis of new organic molecules and measurement of their binding affinities to a protein. Experimental drug research is very time-consuming, risky and comes at an immense cost, usually several billion dollars per drug. Although decades of experimental drug discovery have provided cures for many diseases, there are still diseases for which no effective drug is available. It would be of great benefit to humanity if drug discovery could be performed with theoretical and computational methods and it would likely accelerate the discovery of new drugs. The improved performance of computers and advances in 3D-modelling of proteins during the latest two decades have enabled the development of drugs using theoretical methods. This thesis mainly deals with how quantum mechanics can be used to improve calculated binding free energies, conventionally obtained by simulating proteins and drug molecules according to classical mechanics and with empirical potential-energy functions. This is partly based on methods to combine quantum mechanics and molecular mechanics, originally developed by the Nobel laureate Arieh Warshel.





Printed by Media-Tryck, Lund 2018 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Lund University
Faculty of Science
Division of Theoretical Chemistry
ISBN 978-91-7422-577-8