# Free Energies of Binding from Large-Scale First-Principles Quantum Mechanical Calculations: Application to Ligand Hydration Energies

Stephen J. Fox,[†] Chris Pittock,[†] Christofer S. Tautermann,[‡] Thomas Fox,[‡] Clara Christ,[‡]
N. O. J. Malcolm,[§] Jonathan W. Essex,[†] and Chris-Kriton Skylaris*,[†]

[†]School of Chemistry, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom
[‡]Department for Lead Identification and Optimization Support, Boehringer Ingelheim Pharma GmbH & Co KG, 88397 Biberach, Germany
[§]Accelrys Ltd., 334 Cambridge Science Park, Cambridge CB4 0WN, United Kingdom
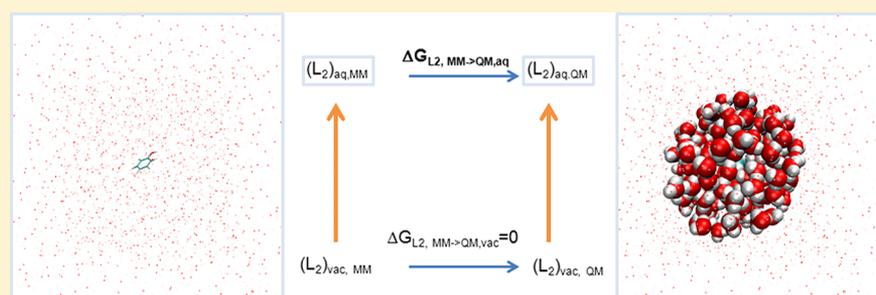
ABSTRACT: Schemes of increasing sophistication for obtaining free energies of binding have been developed over the years, where configurational sampling is used to include the all-important entropic contributions to the free energies. However, the quality of the results will also depend on the accuracy with which the intermolecular interactions are computed at each molecular configuration. In this context, the energy change associated with the rearrangement of electrons (electronic polarization and charge transfer) upon binding is a very important effect. Classical molecular mechanics force fields do not take this effect into account explicitly, and polarizable force fields and semiempirical quantum or hybrid quantum–classical (QM/MM) calculations are increasingly employed (at higher computational cost) to compute intermolecular interactions in free-energy schemes. In this work, we investigate the use of large-scale quantum mechanical calculations from first-principles as a way of fully taking into account electronic effects in free-energy calculations. We employ a one-step free-energy perturbation (FEP) scheme from a molecular mechanical (MM) potential to a quantum mechanical (QM) potential as a correction to thermodynamic integration calculations within the MM potential. We use this approach to calculate relative free energies of hydration of small aromatic molecules. Our quantum calculations are performed on multiple configurations from classical molecular dynamics simulations. The quantum energy of each configuration is obtained from density functional theory calculations with a near-complete psinc basis set on over 600 atoms using the ONETEP program.

1. INTRODUCTION

Rigorous free-energy calculation approaches can provide, in principle, exact values for the relative free energies of binding of a set of ligands to a host. Free-energy perturbation (FEP), thermodynamic integration, and other free-energy methods[1−5] have been in use and in continuous development over the years, motivated by important applications such as drug optimization, where the accurate calculation of free energies is crucial.

High-level ab initio quantum mechanical calculations should ideally be used to describe the receptor–ligand systems in the simulations in order to obtain an accurate and unbiased description of all of the interactions. However, due to the large number of atoms typically involved in free-energy calculations and the large amount of configurational sampling required, ab initio molecular dynamics simulations are typically not computationally feasible, and classical molecular dynamics simulations with force fields are most commonly used. The accuracy of the computed free energy will be affected if the force field Hamiltonian does not reproduce well the intermolecular interactions. For example, polarization of the ligand by the environment (protein or solvent) and the back-polarization of the environment by the ligand are not normally considered explicitly as nonpolarisable force fields are used in most studies. Even in the case of polarizable force field approaches, the error can be significant due to imperfections in the polarization model or the quality of the parametrization.[6] Polarization of a solute by the solvent is often implicitly considered by using solute charges derived by a method that overestimates the solute gas-phase charges to some extent, for

example, RESP 6-31G* charges in the case of the AMBER 94/99/GAFF series of force fields.

The next step in the hierarchy of methods that explicitly include polarization is to employ a hybrid quantum mechanical (QM)−molecular mechanical (MM) or QM/MM approach, where a small part of the system (typically comprised of the ligand) is described by a QM Hamiltonian and the rest of the system by a MM Hamiltonian, and the two descriptions are coupled in a physically suitable way.[7] However, even with QM/MM, performing direct molecular dynamics or Monte Carlo simulations for the lengths of configurational sampling required for free-energy calculations can be prohibitively expensive. To overcome this limitation, approaches have been developed that use a fast and relatively inexpensive Hamiltonian to sample the configurational space and then, on selected configurations, evaluate the energy with the more costly but more accurate QM/MM Hamiltonian.[8−12] The reference Hamiltonian (often a pure MM force field) is first used to estimate the free energy, and this estimate is then corrected by calculating the free energies necessary to change the reference MM Hamiltonian to the QM/MM Hamiltonian. The free-energy change from the MM to QM/MM Hamiltonian is computed via a single-step FEP (Zwanzig equation), using only the structural ensemble generated from the reference MM Hamiltonian. This approach will converge quickly only if the configurational spaces from the MM and QM/MM ensembles have good overlap. However, in practice, this is not guaranteed, and it cannot be predicted. In an effort to improve the overlap of configuration spaces, Beierlein et al.[13] proposed a simpler form of the Zwanzig equation in which only the Coulomb energy is "mutated" from a classical to a QM/MM description. Their approach was formulated to take into account the polarization of the ligand by the host but not the back-polarization of the host by the ligand because only the ligand was described by a QM Hamiltonian. The aim of their work was to obtain reproducible, converged free energies, without the need for closely coupled MM and QM programs. Relative binding free energies were obtained using classical TI, and the Coulomb terms were corrected using QM/MM single-point energy calculations on the MM sampled phase space.

To improve the sampling of the QM/MM ensemble, Woods et al.[14] proposed an approach that uses a Metropolis−Hastings criterion to ensure that the correct configurational distribution is generated for the QM/MM ensemble by accepting only suitable MM configurations. Provided that the simulations are run long enough for sufficient QM/MM configuration space to be sampled, this approach can be used in a FEP context to compute exactly the change in free energy when going from the MM to the QM/MM description.

With the availability of programs capable of performing large-scale QM calculations,[15−18] based mainly on linear scaling density functional theory (DFT) approaches,[19,20] schemes such as the above for QM/MM corrections to free energies have the potential of including in the QM description far larger portions of the system or even the entire system (therefore, a full QM rather than QM/MM description), thus fully taking account of electronic polarization and dispensing with the ambiguities of interfacing a small QM region with an MM region. Toward this goal, we explore here such an approach, and we apply it to calculate the relative free energies of hydration of small aromatic ligands (presented in Figure 1). The water solvent is treated explicitly, and a large number of the water molecules surrounding each ligand is contained within our QM
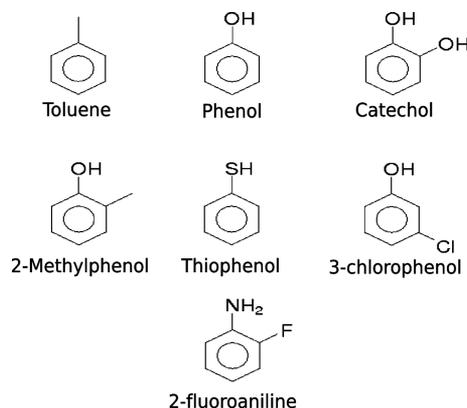

Figure 1. Ligands used in this study.

description. The theoretical and methodological details as well as the setup of the simulations are described in section 2. Our results are presented and discussed in section 3, and we finish with our conclusions in section 4.

2. METHODS

2.1. ONETEP Program. The ONETEP[16] program is a linear scaling DFT code that has been developed for use on parallel computers.[21] ONETEP combines linear scaling with accuracy comparable to conventional cubic scaling plane wave methods, which provide an unbiased and systematically improvable approach to DFT calculations. Its novel and highly efficient algorithms allow calculations on systems containing tens of thousands of atoms.[22] ONETEP is based on a reformulation of DFT in terms of the one-particle density matrix. The density matrix in terms of Kohn−Sham orbitals is

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{n=0}^{\infty} f_n \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}') \tag{1}$$

where f_n is the occupancy and $\psi_n(\mathbf{r})$ are the Kohn−Sham orbitals. In ONETEP, the density matrix is represented as

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha} \sum_{\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta^*(\mathbf{r}') \tag{2}$$

where $\phi_\alpha(\mathbf{r})$ are localized nonorthogonal generalized Wannier functions[23] (NGWFs) and $K^{\alpha\beta}$, which is called the density kernel, is the representation of f_n in the duals of these functions. Linear scaling is achieved by truncation of the density matrix, which decays exponentially for materials with a band gap, and enforcing strict localization of the NGWFs onto atomic regions. In ONETEP, as well as optimizing the density kernel, the NGWFs are also optimized, subject to a localization constraint. Optimizing the NGWFs in situ allows for a minimum number of NGWFs to be used while still achieving plane wave accuracy. The NGWFs are expanded in a basis set of periodic sinc (psinc) functions,[24] which are equivalent to a plane wave basis as they are related by a unitary transformation. Using a plane wave basis set allows the accuracy to be improved by changing a single parameter, equivalent to the energy cutoff in conventional plane wave DFT codes. The psinc basis set provides a uniform description of space, meaning that ONETEP does not suffer from basis set superposition error.[25]

2.2. MM Simulation Setup. We have performed simulations of the small aromatic ligands shown in Figure 1 in water. For the setup of the MD simulations, we started from a catechol molecule (generated in the MOE program[26] and

geometry optimized in the gas phase with the MMFF94x force field[27]) and placed it in a water box containing 1545 explicit waters in a cubic box with periodic boundary conditions in the AMBER version 10[28] package. The TIP3P model[29] was used to describe the water solvent, and the generalized AMBER force field (GAFF)[30] was used to describe the solutes. The AM1-BCC method was used to obtain partial charges for the solutes with the antechamber tool in the AMBER package.

To equilibrate the system, the positions of the hydrogens were relaxed before heating the system from 100 to 300 K with the *NVT* ensemble over 200 ps with positional restraints of 1000 kcal/mol Å^2 on the catechol molecule. Then, we switched to the *NPT* ensemble for 200 ps, keeping the positional restraints on catechol. The system was then run for a further 200 ps with no restraints in the *NVT* ensemble and again switched to *NPT* for 200 ps at 300 K in order to add a final adjustment to the volume of the simulation cell and consequently the density of the water. The simulation cell was constrained to remain cubic, and its final volume had sides of 36.222 Å.

At this point, it was confirmed that the water density, kinetic energy, and potential energy had only small oscillations around a constant value; therefore, the system was deemed to be equilibrated. The catechol molecule was manually mutated in the MOE program to the six other solutes in Figure 1. Production MD simulations were started from the catechol and these new structures (with randomly assigned initial velocities), each containing a ligand in a water box, and run in the *NVT* ensemble at 300 K for 20 ns. Snapshots were taken from the last 18 ns of the trajectory, treating the first 2 ns as further equilibration.

For thermodynamical consistency, we have ensured that the same number of water molecules (1545) was used in all simulations. Furthermore, a cubic simulation cell of length 36.222 Å was used in all simulations to ensure identical basis sets for all subsequent ONETEP calculations. For our MD simulations under the *NVT* ensemble, we used the Langevin thermostat[31] with a collision frequency of 3.0 ps^{-1}, the particle mesh Ewald (PME) algorithm for the long-range electrostatics, a nonbonded cutoff of 8.0 Å, and a time step of 2 fs with the SHAKE algorithm[32] with a tolerance of 1.0×10^{-5} Å.

2.3. MM and DFT Single-Point Energy Calculation Setup. For the MM single-point energy calculations, the GAFF force field was used again but with a nonbonded cutoff of 13.0 Å in a cubic periodic box of length 36.222 Å. Periodic boundary conditions were used with the standard Ewald approach to accurately calculate the electrostatic interactions.

To reduce the computational time for the DFT calculations, we employed an electrostatic embedding (EE) approach for part of the system as implemented in ONETEP.[33] Each QM region was defined as the ligand (solute) and the closest 200 waters (roughly equivalent to a 9.0 Å solvation shell). All of the remaining water molecules were treated as classical embedding charges, where each classical atom center is represented by a tight Gaussian charge distribution. We will thus denote our DFT calculations from now on as "quantum mechanical with electrostatic embedding" (QM EE). The charge given to the classical oxygens was −0.834 e, and that for the classical hydrogens was 0.417 e, as in the TIP3P model. NGWF radii of 4.2 Å were used for all atoms, with four NGWFs on carbon, oxygen, and nitrogen, nine NGWFs on sulfur, fluorine, and chlorine, and one NGWF on hydrogen. A kinetic energy cutoff of 800 eV was used along with the PBE exchange−correlation

functional.[34] Commonly used Kohn−Sham DFT exchange correlation functionals do not account for dispersion, which is crucial for the correct description of the interactions. Within ONETEP, an empirical correction utilizing a damped London potential is used in a DFT+D approach.[35−37] There is a growing number of functionals that explicitly describe dispersion, such as the vdW-DF functionals.[38−40] Also, wave-function-based approaches, such as MP2 or CCSD(T), account for dispersion interactions. However, both approaches are considerably more computationally expensive than the DFT+D approach. We are confident that our calculation parameters produce interaction energies that are very well converged with respect to the basis set as we have previously validated these by comparison to Gaussian basis set calculations with very large basis sets.[33] All simulation cells were cubic and had identical sizes, with a length of 36.222 Å as in the MM calculations, and again, periodic boundary conditions were used with the standard Ewald approach for the electrostatic interactions.[41]

2.4. TI Calculations. TI calculations were performed with the AMBER program. Ligand starting geometries were taken from the starting geometries for the MD simulations. Perturbations were in the direction phenol → new ligand. A two-step method was used, first removing the charges before mutating atoms using softcore potentials[42] for the VdW and then adding the new partial charges. Thirty-nine evenly spaced λ windows were used (with a λ step of 0.025). Each λ step involved the relaxation of the entire system over 100 steps, an equilibration step during which the temperature increased linearly from 100 to 300 K in the *NVT* ensemble over 50 ps, and finally a 200 ps production step in the *NPT* ensemble at 300 K.

Convergence tests were performed using 9 λ windows, 19 λ windowns, and then finally 39 λ windows. The difference in the calculated $\Delta\Delta G$ using 9, 19, or 39 windows is very small. For example, the difference between using 19 windows or 39 windows to calculate the $\Delta\Delta G$ for the phenol → catechol mutation is 0.34 kcal/mol. The error between the forward and reverse calculations reduces from 0.04 kcal/mol for 19 windows to 0.01 kcal/mol when using 39 windows. The maximum difference between 9 and 39 windows is for the phenol → 2-methylphenol mutation with 0.66 kcal/mol, which is reduced to 0.09 kcal/mol when using 19 windows. We have chosen to use 39 windows to ensure tight convergence of the calculated energies.

2.5. MM to QM FEP. We propose that instead of using total energies or just correcting the Coulomb energies as in the approach by Beierlein et al.,[13] we use the interaction energies, defined as

$$\Delta E = E_{S-L} - E_S - E_L \qquad (3)$$

where E_S is the energy of the solvent, E_L is the energy of the ligand, and E_{S-L} is the energy of solvent−ligand complex, following the approach described in ref 33. The MM to QM free-energy differences will now be calculated using

$$\Delta G_{MM \to QM} = -k_B T \ln \langle e^{-(\Delta E^{QM} - \Delta E^{MM})/k_B T} \rangle_{MM} \qquad (4)$$

where ΔE^{QM} is the interaction energy in the quantum description (which includes also the dispersion with the DFT+D approach) and ΔE^{MM} is the interaction energy in the force field description. The notation $\langle \cdots \rangle_{MM}$ denotes an ensemble average over the structures obtained from the MD simulations with the MM force field.
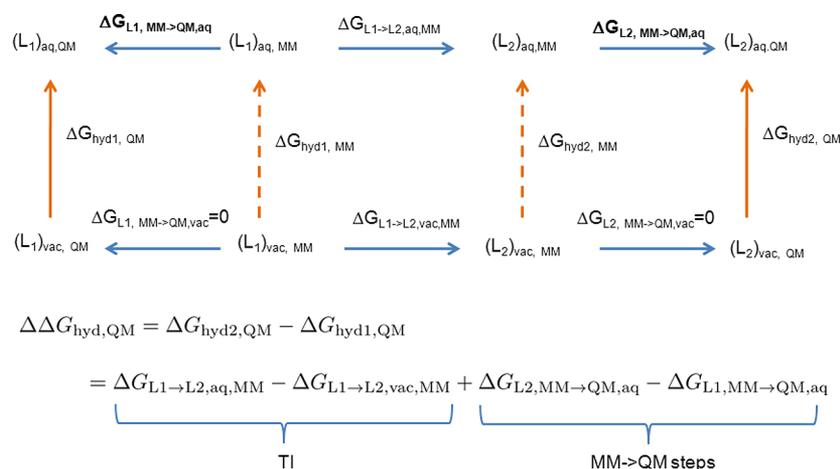
Figure 2. MM → QM-corrected solute hydration free-energy cycle. Going from L_1 to L_2, both are described by QM via a MM alchemical mutation from left to right. The top line depicts the molecule in the solvent $(L)_{aq}$, and the bottom line depicts the molecule in vacuum $(L)_{vac}$.

While formally in eq 4 the total energies E_{S-L}^{QM} and E_{S-L}^{MM} should be used, we have found that such approaches do not work well in practice. One of the issues is that the difference between E_{S-L}^{QM} and E_{S-L}^{MM} is huge, making the direct calculation of the exponentials in eq 4 impossible, but we can overcome this with suitable numerical techniques.[43] However, the biggest issue is that using total energies does not produce converged results to the level in which we are interested, and there is large variation in the obtained free energies within the number of snapshots that we can realistically sample with QM calculations. On the other hand, using interaction energies produces meaningful results as it is more suitable for the problem in question (changes in binding). Interaction energies reduce systematic errors (such as, for example, mismatches between the intramolecular structure between the MM and QM descriptions) and are also very close between the two descriptions (on the order of a few kcal/mol), indicating a better overlap between the MM and QM "interaction phase spaces". The extended thermodynamic cycles needed for the QM-corrected calculation of relative hydration free energies are shown in Figure 2. Because interaction energies are being used and not total energies, changes in internal energies in going from the MM to the QM description for the solutes in vacuum will be taken to be zero; therefore, the $\Delta G_{L1,MM \rightarrow QM,vac}$ and $\Delta G_{L2,MM \rightarrow QM,vac}$ terms are not included in our calculations. The use of interaction energies, instead of total energies, is often the preferred approach;[10,13,44−46] however, it does introduce several approximations. These are the neglect of changes in the solute's and solvent's internal free energy between the MM and QM descriptions. Effectively, this introduces the assumption that the change in internal free energy of a solute in going from MM to QM would be the same in both the vacuum and solvent and would thus cancel each other when calculating relative hydration free energies. We can assume that this is a reasonable approximation, although the degree at which it holds is system-dependent (for example, we would expect a molecule such as catechol to undergo conformational changes upon solvation, which would introduce errors in this assumption). To quantify the error that is introduced by using interaction energies, we would need to see how the obtained free energies differ from the converged free energies obtained using the total energies approach, but this is not computationally feasible.

Shaw et al.[47] examined the compatibility of water models between MM and QM/MM simulations and found that TIP3P and TIP4P both perform well while TIP5P is not at all suitable. For efficient sampling, the overlap of the "phase space" is crucial. In this study, we have used the well-established general AMBER force field (GAFF). Previous work has indicated that there is good conformational overlap between structures generated with GAFF and ONETEP.[48] The assumption of good overlap in this work is probably an acceptable approximation due to the small size and rigidity of the molecules involved.

A one-step FEP based on the Zwanzig formula shown in eq 4 was used to calculate the free-energy change from the MM thermodynamic state of the system to the QM EE state. Snapshots were taken at equidistant time intervals from the last 18 ns of the production trajectories, and the interaction energies were computed with MM and QM EE. In addition to using eq 4, histograms of the interaction energy differences were fitted to different functional forms in order to try to minimize the errors due to finite sampling. We do not know a priori the functional form of these distributions, but we have attempted to use the "guesses" suggested by Nanda et al.[49] One of these choices is a Gaussian form, the results of which we include in this paper. An example of the histogram of the interaction energy differences for phenol is given in Figure 3 along with the fitted Gaussian to the data.

2.6. Calculation Workflow. To perform the calculations needed for the extended thermodynamic cycle scheme of Figure 2 with the above-mentioned techniques, the workflow of calculations shown in Figure 4 was followed. All calculations started from a MD equilibration of a single molecule L1 ("Equilibration" in the figure), out of which the final snapshot was extracted and mutated into the other molecules ("Snapshot" in the figure). From each of these snapshots, TI calculations were run between the solutes ("TI" in the figure) and also separate MD simulations ("MD" in the figure) because the dual topology TI in AMBER does not explicitly simulate the $\lambda = 0$ and 1 points. Finally, snapshots were extracted from the MD simulations, and their interaction energies were computed with the MM and QM EE approaches.
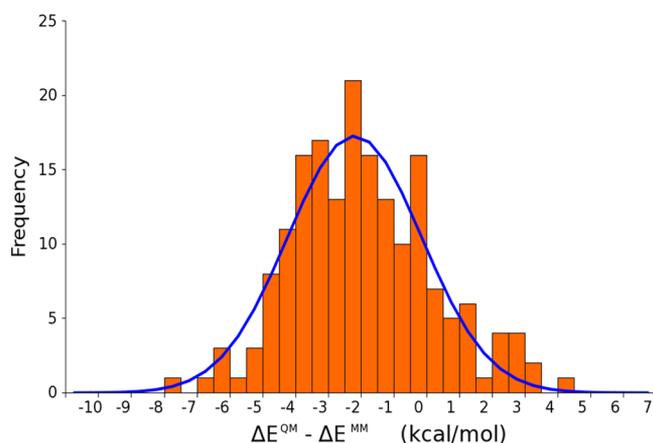
Figure 3. Histogram of interaction energy differences for phenol from the calculation using 180 snapshots. The Gaussian function fitted to the histogram is also shown.
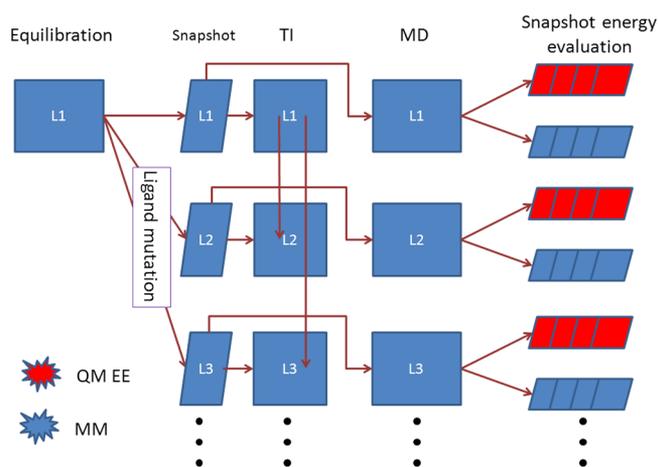

Figure 4. Workflow of calculations. All calculations started from a MD equilibration of a single solute L1, out of which the final snapshot was extracted and mutated into the other solutes. From each of these snapshots, TI calculations were run between the solutes and also separate MD simulations. Finally, snapshots were extracted from the MD simulations, and their interaction energies were computed with the MM and QM EE approaches.

3. RESULTS AND DISCUSSION

The calculated relative hydration free energies are presented in Table 1. All calculated hydration free energies in the table are presented with respect to the phenol molecule whose hydration free energy has been set equal to the experimental value of

-6.62 kcal/mol. Two sets of results are presented, calculations using the set of 90 "odd" snapshots as described in Figure 6 and calculations using the full set of 180 snapshots. Relative hydration free energies ("TI" in Figure 2) as obtained directly from TI with the MM potential are shown in the column "AMBER-TI". The QM-corrected free energies ("$\Delta\Delta G_{hyd,QM}$" in Figure 2) computed with the Zwanzig formula in eq 4 are shown in the column "Zwanzig". We also present the QM-corrected free energies, which have been obtained from the Gaussian fits to the probability distributions (as in Figure 3) instead of using eq 4, in the column "Gauss". In Table 1, we also show the relative hydration free energies obtained by direct calculation of the difference $\Delta G_{hyd2,QM} - \Delta G_{hyd1,QM}$ in Figure 2 using the minimal parameter implicit solvation model of ONETEP,[50] shown in column "ONETEP IS". These implicit solvent calculations were performed only on the set of 90 odd snapshots as the obtained hydration free energies are simple averages of the solvent model hydration free energy for each snapshot and converge rather rapidly and predictably (the convergence error in these values is less than 0.1 kcal/mol). As reference values "SMD (exp)", we use the hydration free energy obtained from implicit solvation calculations with the SMD model.[51] These calculations were performed with the Gaussian09[52] program at the M052x/6-31G(d) SCRF-(IEFPCM,Solvent=Water,SMD) level on a single geometry-optimized structure. Experimentally determined hydration free energies are available for phenol,[53] toluene,[53] thiophenol,[53] catechol,[54] 2-methylphenol,[55] and 3-chlorophenol[56] and are included in brackets, while for 2-fluoroaniline, we do not have experimental data available. The errors for these calculations are estimated in Table 2 and discussed later.

For small mutations in the solutes, as with the molecules in this study, the standard TI approach predicts the relative hydration free energies well and has a very good correlation with experiment, with an R^2 value of 0.98. The largest error seen for the MM TI results is 1.4 kcal/mol (for toluene) with a mean unsigned error of 0.9 kcal/mol and a rms error of 0.9 kcal/mol. As TI is among the most theoretically rigorous free-energy approaches for including the entropic component of the free-energy change, the good results in Table 1 (under the assumption that entropy changes are captured correctly by the TI) would imply that the enthalpic component is well described for these molecules, or in other words, for these cases, the force field describes the solute–solvent interactions well.

For the 90 and 180 snapshots results with the Zwanzig equation, the trend that we observe is that when the AMBER TI values deviate by more than $\sim k_BT$ (about 0.6 kcal/mol) from the experimental value, QM offers an improvement, while

Table 1. Hydration Free Energies Relative to the Phenol Experimental Hydration Energy Using 90 and 180 Snapshots in the MM to QM FEP[a]

ligand	Zwanzig-90	Zwanzig-180	Gauss-90	Gauss-180	AMBER-TI	ONETEP IS[b]	SMD (exp)
catechol	-9.34 ± 1.26	-9.63 ± 0.62	-12.66	-11.60	-8.66	-9.73	$-9.3\ (-9.4)$
toluene	-0.84 ± 1.05	-0.97 ± 0.55	-1.45	-0.68	-2.29	-1.46	$-1.3\ (-0.9)$
3-chlorophenol	-6.59 ± 1.03	-7.32 ± 0.69	-8.00	-7.25	-6.77	-6.53	$-6.7\ (-6.6)$
2-fluoroaniline	-5.05 ± 1.05	-5.10 ± 0.58	-4.84	-4.66	-5.83	-6.04	$-4.5\ (--)$
2-methylphenol	-6.50 ± 1.06	-6.81 ± 0.57	-9.00	-7.80	-6.52	-6.04	$-6.3\ (-5.9)$
thiophenol	-4.86 ± 1.45	-4.75 ± 1.02	-11.34	-6.12	-3.39	-2.84	$-2.3\ (-2.6)$
phenol (ref)	-6.62 ± 0.67	-6.62 ± 0.36	-6.62	-6.62	-6.62	-6.62	$-6.6\ (-6.6)$

[a]The errors in the "Zwanzig-90" and "Zwanzig-180" columns are calculated from bootstrapping the data. Comparisons are made to hydration energies from Gaussian SMD calculations and experimental values where available. Energies are given in kcal/mol. [b]Using only 90 snapshots.

Table 2. Standard Deviation (in kcal/mol) of the MM → QM Correction (eq 4) for Each Solute, Calculated in Three Ways[a]

	catechol	3-chlorophenol	2-fluoroaniline	2-methylphenol	phenol	thiophenol	toluene
Gaussian fit	0.64	0.35	0.17	0.69	0.36	2.72	0.20
Zwanzig	0.11	0.28	0.24	0.02	0.39	1.70	0.22
50%	0.15	0.44	0.31	0.15	0.41	0.92	0.23
66%	0.21	0.14	0.18	0.18	0.37	0.47	0.11
90%	0.08	0.16	0.05	0.05	0.19	0.04	0.02
bootstrapping	0.26	0.33	0.22	0.21	0.36	0.66	0.19

[a](1) Generated from four sets of 90 snapshots (50% of data points), as shown in Figure 6. This was done using the Zwanzig equation and the Gaussian fit. (2) Data generated by taking a random 50, 66, or 90% of the original 180 snapshots, 10 times, on only the Zwanzig data. (3) Bootstrapping, taking 180 snapshots with replacements from the original 180 snapshots, 100 times, on only the Zwanzig data.

when the MM TI result is already very good, the QM correction does not make it worse (we should note that for 3-chlorophenol and 2-methylphenol, the QM-corrected results appear worse than the MM TI, but they are within the estimated error of the experimental result). There is a notable exception to this trend for the case of thiophenol, whose value is made much worse when the QM correction is applied. The source of this error is hard to pinpoint; it could be due to an inadequate description of the interactions of thiophenol with the chosen exchange−correlation functional or the inadequate configurational sampling, which is discussed later. However, even with this exception, it is remarkable that we get a notable improvement for the rest of the solutes, which should be due to the fact that the QM EE approach includes the full electronic charge transfer and polarization, which cannot be described explicitly with the force field. The results from the Zwanzig equation using 180 snapshots agree less with the experimental values; however, the trends that we noticed with the 90 snapshots remain, and the less good agreement has to be attributed to the errors in the convergence of the exponential averaging in the Zwanzig equation, which implies that at 90 snapshots, the Zwanzig equation has not converged. This is supported by the fact the bootstrap errors are almost twice as large as those for 180 snapshots. The reduced error for 180 is promising, but we see that it is still not converged. Given more computational resources, a greater number of snapshots could be obtained to investigate convergence more thoroughly.

The Gaussian fit values show no improvement over standard TI using 90 or 180 snapshots, simply indicating that a Gaussian function is not suitable for representing these distributions. The ONETEP implicit solvation results appear to be surprisingly accurate and substantially better for thiophenol, reducing its error by an order of magnitude to 0.2 kcal/mol. This better result can be attributed to the fact that in the ONETEP IS calculation, we consider the averaged interaction of the molecule with the continuum describing the solvent. In this case, the interaction terms (electrostatic, dispersion, and entropic) have less direct dependence on the exchange−correlation functional than those in the Zwanzig interaction energy approach, which is based on the interaction energies with explicit water molecules.

The convergence of the MM to QM correction of eq 4 as a function of increasing the number of snapshots in the average is an important point that we need to consider, as shown in Figure 5. We observe the well-known "sawtooth"[57] behavior due to the fact that certain snapshots can have a significant influence on the free-energy value. Even though the energy appears to be converging, with a variation of less than 0.5 kcal/mol during the last 10 snapshots, we have no guarantee that there is not going to be another large drop if we are able to
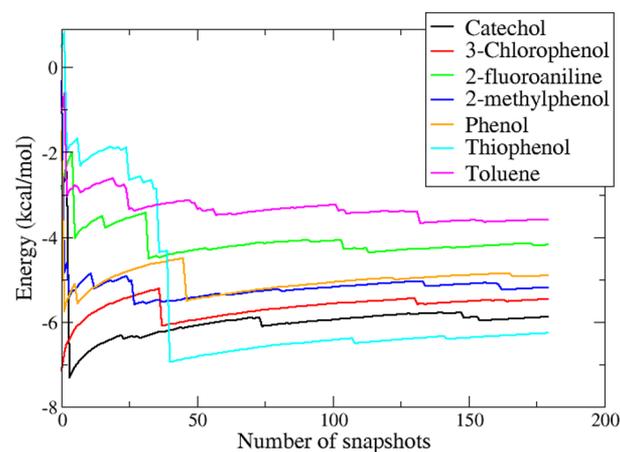

Figure 5. Convergence of the MM → QM free-energy change as a function of the number of snapshots included in the average.

increase our sampling to a larger number of snapshots. This is a consequence of the exponential form of the Zwanzig equation, where the tails of the energy distribution dominate the average and produce unpredictable convergence behavior.

To obtain estimates of the degree of convergence of our MM → QM free-energy corrections (eq 4), the standard deviation for each solute was calculated by three different approaches, as presented in Table 2. The first approach consists of splitting up the 180 snapshots into four sets of 90 snapshots, as shown in Figure 6. The second approach selects 50, 66, or 90% of the snapshots 10 times and calculates the standard deviation from the $\Delta G_{MM \rightarrow QM}$ values obtained. Finally, we have also performed resampling by bootstrapping the data.[58] This approach takes 180 snapshots with replacements (i.e., a particular snapshot can occur more than once in each resampled set of 180) and calculates the standard deviation of 100 such resamples. For the Gaussian fits, only the first approach was used.

For results obtained using eq 4, we observe that the standard deviation is ~0.4 kcal/mol or less for each molecule, except thiophenol. This suggests that our free-energy results for thiophenol are more likely to contain larger errors than the rest of the solutes, and more extensive sampling would ideally be needed for this molecule. The values in Table 2 can give us an indication of the sensitivity of the free-energy corrections to the selection and number of snapshots. Although they cannot guarantee that there will not be any more abrupt jumps in the free energies as in Figure 5 were we to increase our sample size to more than 180 snapshots, they do give us an indication of the variability of free energies calculated for sizes of samples of the order that we have used here. For example, we have observed that the hydration free energies that are obtained by
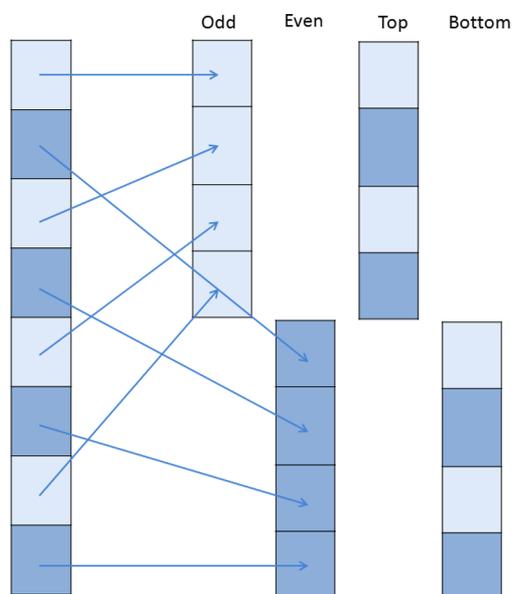
Figure 6. Schematic showing the four different ways that the 180 snapshots were divided into sets of 90 to estimate errors.

using the averages of the resampling approaches 2 or 3 differ from those of the "Zwanzig-180" column of Table 1 by less than 0.1 kcal/mol, which is encouraging. With the Gaussian fit approach, generally larger deviations are observed in Table 2 for the solutes, being up to 0.7 kcal/mol, except again for thiophenol, which has a much larger error of 2.7 kcal/mol.

4. CONCLUSIONS

In this study, we have used an extended free-energy cycle with the aim of obtaining relative hydration free energies from large-scale DFT calculations (with more than 600 atoms) within a theoretical framework (one-step FEP) that at least in principle fully accounts for the entropic contributions to free energy. Our aim with this approach was to add a correction to the relative free energies obtained from a standard TI procedure with a classical force field in such a way as to obtain the free energies that would have been obtained if the TI simulations had been performed throughout by DFT calculations. Using DFT has the obvious advantage that the electronic degrees of freedom are explicitly included in the calculation and thus fully account for the energetic contribution of the electronic charge transfer and polarization that takes place in molecular association events.

Although an improvement of free energies calculated with our approach toward the experimentally measured values would be ideal, this may not always be possible. The results from the Zwanzig equation converge, in the limit of a large number of snapshots, to the QM hydration free energy. Even if no other sources of error are encountered, the obtained relative free energies would be subject to errors in the description provided by DFT, such as the quality of the exchange–correlation functional, the pseudopotential, and the basis set (although, in this case, we are fairly confident that we have used a large, very accurate basis set). We have applied our approach to calculate the relative hydration free energies of small aromatic molecules, and we have found that in most cases, we obtain improvements over the values produced with the classical force field. However, the improvements are not universal, as in the case of thiophenol, and our analysis shows a much larger deviation. This indicates that the number of snapshots sampled is not

enough or that there is poor overlap between the MM and QM EE Hamiltonians. A future extension to this work would be to explore the compatibility of the QM and MM description, with modifications of the force field to improve the overlap between the descriptions.

The free-energy calculations reported here rely on the use of the Zwanzig equation, which, due to the exponential averaging, is notoriously difficult to converge and can be heavily affected by rare configurations. Even though we see a degree of convergence to well within chemical accuracy (less than 0.5 kcal/mol) with respect to the number of snapshots sampled for our largest sets of 180 configurations, there is no guarantee that additional sampling may not yet access a new rare configuration that could cause a substantial change in the calculated free energies. We should also keep in mind that we have used here rather rigid molecules; therefore, for larger and more flexible species, the number of configurations that will need to be included in the sampling process will need to be larger.

On the other hand, this process is "trivially parallel"; therefore, given sufficient computational resources, a significantly larger number of configurations would be possible to sample with DFT calculations. Our DFT calculations were performed with the ONETEP code, which is already capable of doing large-scale DFT calculations on entire protein ligand complexes; however, for this method to be used for the computation of relative protein–ligand free energies of binding, more research is needed. The next step would be to apply this approach to model binding sites and to explore both the effect of significantly increasing the number of configurations evaluated with the DFT approach as well as investigating the compatibility of the exchange–correlation functional with the classical potential used. We are currently investigating such model systems toward developing large-scale DFT calculation methods that can be used for the final stages of refinement in computational drug optimization.

■ AUTHOR INFORMATION

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
(2) Gilson, M. K.; Zhou, H.-X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
(3) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2010**, *31*, 1569.
(4) Gallicchio, E.; Levy, R. M. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 27.
(5) Michel, J.; Essex, J. *J. Comput. Aided Mol. Des.* **2010**, *24*, 639.
(6) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Mobley, D. L.; Schnieders, M. J.; Haque, I.; Lambrecht, D. S.; DiStasio,

R. A., Jr.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549.

(7) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

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

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

(10) Liu, W.; Sakane, S.; Wood, R. H.; Doren, D. J. *J. Phys. Chem. A* **2002**, *106*, 1409.

(11) Štrajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333.

(12) Bandyopadhyay, P. *J. Chem. Phys.* **2005**, *122*, 091102.

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

(14) Woods, C. J.; Manby, F. R.; Mullholland, A. J. *J. Phys. Chem.* **2008**, *128*, 014109.

(15) Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez, D. *J. Phys.: Condens. Matter* **2002**, *14*, 2745.

(16) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *J. Chem. Phys.* **2005**, *122*, 084119.

(17) Bowler, D. R.; Miyazaki, T.; Gillan, M. J. *J. Phys.: Condens. Matter* **2002**, *14*, 2781.

(18) Ufimtsev, I. S.; Martinez, T. J. *Comput. Sci. Eng.* **2008**, *10*, 26.

(19) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085.

(20) Bowler, D. R.; Miyazaki, T. *Rep. Prog. Phys.* **2012**, *75*, 035603.

(21) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *Phys. Status Solidi* **2006**, *243*, 973.

(22) Hine, N. D. M.; Haynes, P. D.; Mostofi, A. A.; Skylaris, C.-K.; Payne, M. C. *Comput. Phys. Commun.* **2009**, *180*, 1041.

(23) Skylaris, C.-K.; Mostofi, A. A.; Haynes, P. D.; Diéguez, O.; Payne, M. C. *Phys. Rev. B* **2002**, *66*, 035119.

(24) Mostofi, A. A.; Haynes, P. D.; Skylaris, C.-K.; Payne, M. C. *J. Chem. Phys.* **2003**, *119*, 8842.

(25) Haynes, P. D.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *Chem. Phys. Lett.* **2006**, *422*, 345.

(26) *MOE2009.10*; Chemical Computing Group Inc.: Montreal, 2009.

(27) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.

(28) Case, D. A.; Darden, T.; Cheatham, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Roitberg, S. H. A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D.; Seetin, M.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER10*; University of California, San Francisco, 2008.

(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.

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

(31) Alderman, S. A.; Doll, J. D. *J. Chem. Phys.* **1976**, *64*, 2375.

(32) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(33) Fox, S.; Pittock, C.; Fox, T.; Tautermann, C.; Malcolm, N.; Skylaris, C.-K. *J. Chem. Phys.* **2011**, *135*, 224107.

(34) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(35) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(36) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(37) Hill, Q.; Skylaris, C.-K. *Proc. R. Soc. London, Ser. A* **2009**, *465*, 669.

(38) Dion, M.; Rydberg, H.; Schroder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys Rev. Lett.* **2004**, *92*, 246401.

(39) Dion, M.; Rydberg, H.; Schroder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys Rev. Lett.* **2005**, *95*, 109902.

(40) Lee, K.; Murray, E. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys Rev. B* **2010**, *82*, 081101(R).

(41) Hine, N. D. M.; Dziedzic, J.; Haynes, P. D.; Skylaris, C.-K. *J. Chem. Phys.* **2011**, *135*, 204103.

(42) Case, D. A.; Mobley, D. L.; Steinbrecher, T. *J. Chem. Phys.* **2007**, *127*, 214108.

(43) Berg, B. A. *Comput. Phys. Commun.* **2003**, *153*, 397.

(44) Wesolowski, T.; Warshel, A. *J. Phys. Chem.* **1994**, *98*, 5183.

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

(46) Rod, T. H.; Rydberg, P.; Ryde, U. *J. Chem. Phys.* **2006**, *124*, 174503.

(47) Shaw, K.; Woods, C.; Mullholland, A. *J. Phys. Chem. Lett.* **2010**, *1*, 219−223.

(48) Fox, T.; Thomas, B. E., IV; McCarrick, M.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 10779.

(49) Nanda, H.; Lu, N.; Woolf, T. B. *J. Chem. Phys.* **2005**, *122*, 134110.

(50) Dziedzic, J.; Helal, H. H.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *Europhys. Lett.* **2011**, *95*, 43001.

(51) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B.* **2009**, *113*, 6378.

(52) Frisch, M. J.; Trucks, G.W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian Inc.: Wallingford, CT, 2009.

(53) Marenich, A. V.; Kelly, C. P.; Thompson, J. D.; Hawkins, G. D.; Chambers, C. C.; Giesen, D. J.; Winget, P.; Cramer, C. J.; Truhlar, D. G. *Minnesota Solvation Database*, version 2012; University of Minnesota: Minneapolis, MN, 2012.

(54) Mordasini, T. Z.; McCammon, J. A. *J. Phys. Chem. B* **2000**, *104*, 360.

(55) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2001**, *23*, 517.

(56) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350.

(57) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235.

(58) Ader, H. J.; Mellenbergh, G. J.; Hand, D. J. *Advising on Research Methods: A Consultant'S Companion*; Johannes van Kessel Publishing: Huizen, The Netherlands, 2008.