JCTC Journal of Chemical Theory and Computation

Protein−Ligand Complexes: Computation of the Relative Free Energy of Different Scaffolds and Binding Modes

Julien Michel,[†] Marcel L. Verdonk,[‡] and Jonathan W. Essex*,[†]

School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom, and Astex Therapeutics Ltd., 436 Cambridge Science Park, Cambridge, CB4 0QA, United Kingdom

Received April 2, 2007

Abstract: A methodology for the calculation of the free energy difference between a pair of molecules of arbitrary topology is proposed. The protocol relies on a dual-topology paradigm, a softening of the intermolecular interactions, and a constraint that prevents the perturbed molecules from drifting away from each other at the end states. The equivalence and the performance of the methodology against a single-topology approach are demonstrated on a pair of harmonic oscillators, the calculation of the relative solvation free energy of ethane and methanol, and the relative binding free energy of two congeneric inhibitors of cyclooxygenase 2. The stability of two alternative binding modes of an inhibitor of cyclin-dependent kinase 2 is then investigated. Finally, the relative binding free energy of two structurally different inhibitors of cyclin-dependent kinase 2 is calculated. The proposed methodology allows the study of a range of problems that are beyond the reach of traditional relative free energy calculation protocols and should prove useful in drug design studies.

Introduction

Free energy is an important thermodynamic property, and its knowledge permits the prediction of a wide variety of chemical phenomena ranging from binding to phase transitions.[1] Ever since the first applications of the free energy perturbation methodology (FEP) to the calculation of relative free energies were reported,[2] considerable methodological efforts have been devoted to improving the simulation protocols. These efforts are partly justified by the reports of several studies, confirming that relative binding free energies of protein−ligand complexes can be predicted with good precision and accuracy, making the technology an attractive tool for drug design.[3−7] Together with a tremendous increase in computational power, these refinements have made binding free energy calculations sufficiently rapid such that it becomes feasible in certain circumstances to consider their routine application in a drug design environment.[8] The free energy difference between two molecules A and B in a given medium can be calculated for example by thermodynamic integration:

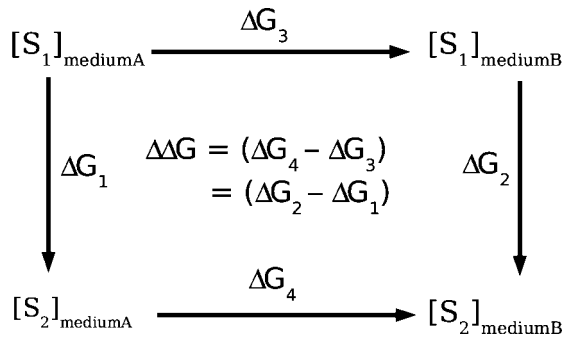
$$\Delta G_{\text{medium,A}\rightarrow\text{B}} = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda}\,\mathrm{d}\lambda = \int_0^1 \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \quad (1)$$

λ is a coupling parameter that allows the smooth transformation of the potential energy function $U(\lambda=0)$, appropriate for molecule A, into a potential energy function appropriate for molecule B, $U(\lambda=1)$. The brackets in eq 1 denote an ensemble average corresponding to the derivative of the potential energy function $U(\lambda)$ with respect to λ (free energy gradients).

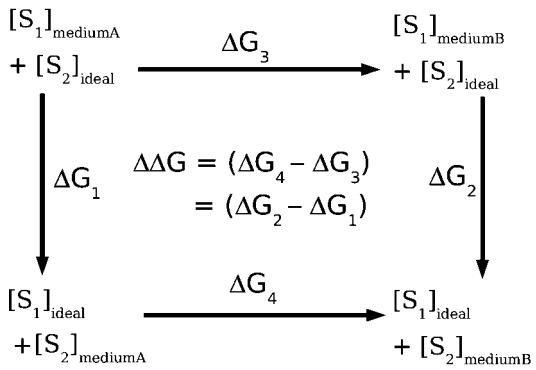
Most protein−ligand binding free energy studies have considered series of congeneric inhibitors of a protein. Two main reasons dictated these choices. First, the free energy perturbation or thermodynamic integration equations converge more easily for similar compounds. This difficulty can be circumvented by running longer simulations. As increasing computer power becomes more and more affordable, this solution becomes increasingly feasible. Second, a practical scheme for the smooth transformation of the potential energy function A $U(\lambda=0)$ into the potential energy function B

* Corresponding author e-mail: J.W.Essex@soton.ac.uk.
[†] School of Chemistry, University of Southampton.
[‡] Astex Therapeutics Ltd.

10.1021/ct700081t CCC: $37.00
Published on Web 08/04/2007



(a) single topology thermodynamic cycle



(b) dual topology thermodynamic cycle

Figure 1. Thermodynamic cycles that relate the difference in free energy between S_1 and S_2 in two media A and B. S_1 and S_2 could be two small molecules and media A and B, water and vacuum, in which case, the double free energy difference will correspond to the relative hydration free energy of S_2 with respect to S_1. If the media A and B represent a solvated protein and pure water, then the double free energy difference will correspond to the relative binding free energy of S_2 with respect to S_1. While the horizontal processes corresponding to ΔG_3 or ΔG_4 are often measured experimentally, the vertical processes corresponding to ΔG_1 or ΔG_2 are usually easier to calculate in a computer simulation. The first cycle implements the single-topology approach where S_1 is converted into S_2 by smooth variation of its force field parameters and geometry. The second cycle implements the dual-topology approach where the interaction energy of S_1 with its medium is gradually turned off while the interaction energy of S_2 is gradually turned on.

$U(\lambda=1)$ has to be proposed. This is often done by a single-topology paradigm shown in Figure 1a. In this approach, the free energy difference between two molecules A and B is calculated by interpolating the force field parameters of molecules A and B. This often requires the coupling of internal coordinate changes to the coupling parameter λ, and the presence of dummy atoms if the topologies of molecules A and B differ. This scheme makes it difficult to devise internal coordinate changes that would convert molecule A into a structurally different molecule B. As a result, many ligands that do not share a similar topology are not considered for a free energy simulation because they would require a complex system setup. This severely limits the applicability of the technique in a rational drug design context. A methodology that readily allows the consideration of sets of structurally diverse ligands is therefore desirable. In the dual-topology paradigm described in Figure 1b, the force field parameters and internal coordinates of molecules A and B are no longer coupled with λ. Rather, a free energy change occurs because the interaction energy of the molecules A and B are coupled to λ.[9] In this approach, compounds A and B can have arbitrary geometries as no scheme to convert the internal coordinates of molecules A into B is necessary, and it seems therefore that the problem of ligand diversity can be solved by this technique. However, practical applications of the dual-topology paradigm suffer from two main difficulties which have limited usage of this technique.

First, noisy free energy gradients can be recorded at the beginning or end of the perturbation ($\lambda = 0.0$ or 1.0). This problem is related to the functional form of the Lennard-Jones (LJ) nonbonded equation, which makes it difficult to turn off completely the LJ interaction at one site. This difficulty can be avoided by not running simulations at the end states and extrapolating the free energy gradients.[10] A more satisfactory solution is to use softened intermolecular energy functions that ensure that the LJ interaction can be turned off smoothly.[11,12] Second, at either end of the simulation, one compound is completely decoupled from its environment. As a result, it could drift away from its initial position, leading to serious convergence issues. A related problem occurs if the intramolecular energy of the ligand is turned off, as in this situation bonds between atoms are broken. If the bond energy is described by a harmonic term, this invariably leads to a divergence of the free energy change.[13] These problems can be avoided by enforcing a restraint that keeps the ligands in the binding site and by not turning off the intramolecular energy terms.[13−16]

Other more original approaches that allow the calculation of binding free energies for diverse compounds have been proposed. Schafer et al. have suggested the use of a nonphysical reference compound that is designed to maximize phase space overlap with a series of compounds of interest. A single simulation is then carried out with this reference compound, and numerous configurations are stored for analysis.[17,18] Each compound in the series can then be mapped onto the reference compound. The free energy difference is then calculated by traditional FEP. The approach is, in principle, very efficient, as a single simulation of a protein−ligand complex is required. However, it can be difficult to devise a reference compound that has a good phase space overlap with all the compounds to be studied, making the calculations potentially imprecise.[19] In addition, ligand flexibility has yet to be addressed.[20]

Another approach that has recently gained popularity involves the calculation of absolute rather than relative binding free energies by either the double decoupling method,[21] potential of mean force approaches,[22] or a combination of these.[23] Absolute binding free energy calculations can be demanding, as the complete annihilation of a ligand can require extensive simulation of several intermediate states to yield precise answers. In addition, it can be difficult to deal with the large structural changes that can occur in the binding site if the ligand is removed. For example, a large

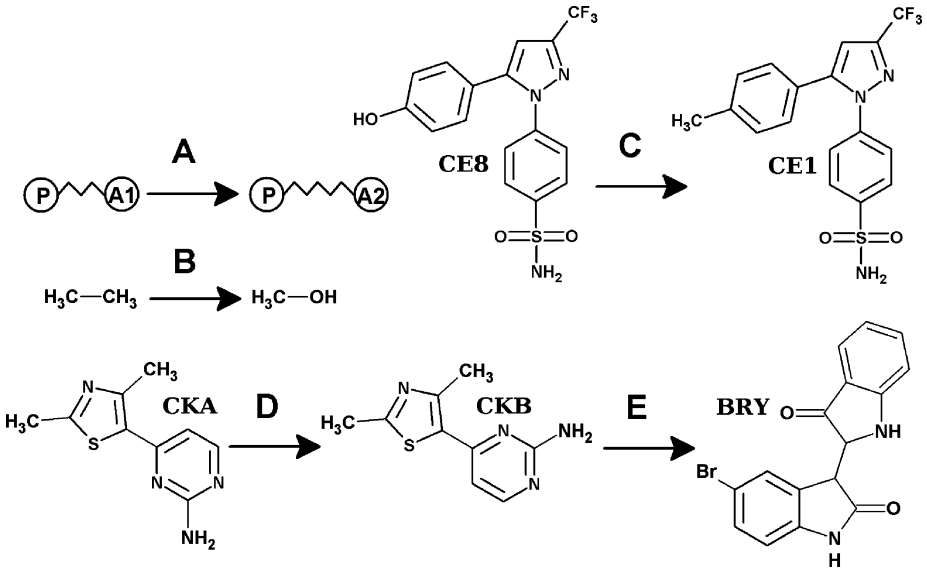


Figure 2. The systems considered in this study. System A is the perturbation of harmonic oscillator P-A1 into harmonic oscillator P-A2. System B is the perturbation of ethane into methanol in a box of TIP4P water. System C is the perturbation between CE8 and CE1, two congeneric inhibitors of the protein COX2. System D is the perturbation of one binding mode of an inhibitor of CDK2, denoted CKA, into an alternative binding mode to the same protein, denoted CKB. System E is the perturbation of the previous inhibitor in the binding mode CKB, into another CDK2 inhibitor, 5-bromoindirubin, and denoted BRY.

change in hydration pattern and conformation of the binding site of the protein OppA is observed between unliganded and liganded structures.[24,25] HIV protease is also known to undergo substantial conformational changes upon inhibitor binding.[26] Such large conformational changes are not sampled easily and rapidly by conventional Monte Carlo or molecular dynamics simulations, leading to potentially imprecise absolute binding free energies.

Methods that calculate relative binding free energies suffer less from these issues, as the binding site is always occupied by a ligand, and it can be expected that relative binding free energies will converge more rapidly than absolute binding free energies. In addition, if the ligands share some common structural features, they often benefit from cancellation of systematic errors in the force field parameters. Finally, the experimental binding affinity data are often available in the form of IC50s, and their conversion to an absolute free energy scale is not always straightforward, making direct comparison of the calculated absolute binding free energies more difficult than their relative counterparts.

In this article, we show that it is possible to efficiently compute the relative binding free energy of substantially different ligands. This is achieved by combining Monte Carlo sampling with a dual-topology approach and a constraint. The formulation is general and demonstrated to give results identical to a single-topology protocol. We then apply the methodology to classes of problems that are not easily handled by a single-topology approach. The range of perturbations covered in this study is shown in Figure 2.

Many existing dual-topology implementations used in the context of relative free energy calculations are in fact best described as hybrid single-/dual-topology methodologies because a portion of the ligand is invariant during the perturbation.[14,16,20] This is necessary to avoid the ligand drifting when it is fully decoupled from its environment, as is commonly experienced in absolute binding free energy calculations.[15] For such a scheme to be practical, the two ligands of interest must share common structural features and occupy the same regions of space. The present work builds on previous efforts from diverse groups[13−16,19−22,27] and strives to overcome these limitations to propose an efficient binding free energy calculation scheme generally applicable to ligand binding studies.

Methods

The free energy calculations were performed with the program ProtoMS2.1.[28] The replica exchange thermodynamic integration (RETI) method[29] was used to construct the free energy profiles, and the necessary ensemble of states were formed using Metropolis Monte Carlo sampling.[30] In the simulations of ethane and methanol, the solvent was represented by a periodic box of 533 TIP4P molecules.[31] In the other systems, the solvent was either modeled by a ball of TIP4P water of 22 Å radius centered on the ligands or by a generalized Born surface area (GBSA) model.[32] The GBSA model employed here is an implementation of the pairwise descreening approximation of Hawkins et al.[33] which has been parametrized to be used in conjunction with AM1/BCC atomic partial charges.[34] Because the generalized Born interaction energy term cannot be broken down into pairwise terms, it does not integrate well with typical Monte Carlo simulations. We have shown, however, that with the adoption of specialized Monte Carlo moves it is possible to sample rigorously the equilibrium distribution of a biomolecular system solvated by an accurate GBSA potential with a crude, more efficient potential and thus recover most of the efficiency loss with minimal approximations.[35]

Models and parameters for the solutes were obtained from a previous study[36] or were derived using the GAFF force field[37] and the AM1/BCC method[38] to obtain atomic partial

charges. Models for the proteins were also taken from a previous study or created according to similar guidelines.[36] The PDB codes of the crystallographic structures used to construct the protein models were 1CX2 (COX2),[39] 1PXJ (CDK2 inactivated),[40] and 2C5O (CDK2 activated).[41]

The bond angles and torsions for the side chains of protein residues within 10 Å of any heavy atom of the ligand and all the bond angles and torsions of the ligand were sampled during the simulation, with the exception of rings. The bond lengths of the protein and ligand were constrained. A 10 Å residue-based cutoff feathered over the last 0.5 Å was employed in all simulations. In the generalized Born simulations, a cutoff of 20 Å for the calculation of the Born radii was applied.

For the explicit solvent simulations of the inhibitors in the bound state, solvent moves were attempted with a probability of 85.7%, protein side-chain moves with a probability of 12.8%, and solute moves with a probability of 1.4%. In the unbound state, solvent moves were attempted 98.4% of the time. In the simulation of ethane and methanol, solvent moves, solute moves, and volume moves were attempted 99%, 0.9%, and 0.1% of the time, respectively. The move probabilities were selected on the basis of the number of protein residues, solvent molecules, and according to previous studies.[35,36] All of the simulations were conducted at 25 °C.

For the perturbation of ethane into methanol, the free energy gradients were accumulated at 11 equally spaced values of the coupling parameter λ. The system was first well pre-equilibrated at one value of the coupling parameter, and each simulation was further equilibrated for 5 million (M) moves before collecting statistics for 25M moves. For the congeneric COX2 inhibitors, 12 values of the coupling parameter λ were employed (0.00, 0.10, ..., 0.90, 0.95, and 1.00) in the single-topology calculations to be consistent with a previous study,[36] and the equilibration and data collection phases consisted of 10M and 30M moves, respectively. For the implicit solvent simulations, these quantities were reduced to 100 000 (100K) and 1M moves, respectively. For the perturbations in CDK2, 21 equally spaced λ values were used; each window was equilibrated for 30 M moves, and data were collected for 50M moves (750K and 1.8M moves, respectively, in the implicit solvent simulations). For a given set of conditions, the free energy change was taken as the mean of five independent simulations and the error estimate as one standard error from the mean. Depending on the system and the simulation conditions, each simulation required 12−36 h on 11−21 2.2 GHz Opteron machines. The calculation of the free energy differences by the dual-topology method was achieved by coupling the two solutes of interest through eq 2:

$$U(\lambda) = U_0 + \lambda U(S_2) + (1 - \lambda)U(S_1) \qquad (2)$$

where $U(S_x)$ represents the energy terms associated with the solute X being turned off or on, and U_0 represents the energy terms for the rest of the system. Note that only the intermolecular energy of the solutes is turned on or off. The intramolecular nonbonded energy and the bonded terms are not modified. A fully decoupled solute is thus transferred to a gas-phase environment. As a result, in the dual-topology method, only intermolecular energy terms contribute to the free energy gradients. By contrast, in the single-topology method, intramolecular as well as intermolecular terms typically contribute to the free energy gradients. To avoid numerical instabilities when one solute is fully decoupled from its surrounding medium, a separation-shifted soft-core functional form (eq 3) for the solute i nonbonded energy was implemented:[12]

$$U_{\text{nonbonded},\lambda} = (1 - \lambda)4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}^{12}}{(\lambda\delta\sigma_{ij} + r_{ij}^{2})^{6}}\right) - \left(\frac{\sigma_{ij}^{6}}{(\lambda\delta\sigma_{ij} + r_{ij}^{2})^{3}}\right)\right] + \frac{(1 - \lambda)^{n} q_i q_j}{4\pi\epsilon_0\sqrt{(\lambda + r_{ij}^{2})}} \qquad (3)$$

where the parameters n and δ were introduced to control the softness of the Coulombic and Lennard-Jones interactions, respectively. These parameters were adjusted for individual perturbations to ensure optimum softening. For the implicit solvent simulations, the generalized Born energy was scaled by the same exponent n, while linear scaling was used for the surface area term. The internal degrees of freedom of each solute were sampled independently. However, to avoid solute drift at either end of the perturbation, the translations and rotations of the pair of ligands were coupled; that is, each individual Monte Carlo move translates and rotates both ligands by the same amount. Internal energy terms are not turned on or off, and a fully decoupled solute is thus in an ideal state. In this state, the free energy of the solute is invariant to rigid body translations and rotations, and the present constraint does not therefore contribute to the free energy change. Such a constraint also has the advantage of being easily implemented in a Monte Carlo simulation package.

In essence, the present dual-topology technique can be considered as a large single-topology perturbation where all the atoms of the first ligand are gradually converted from fully interacting atoms to dummy atoms, while all the atoms of the second ligand are gradually converted from dummy atoms to fully interacting atoms with the additional presence of one dummy bond between the center of geometry of the two ligands. Because the intramolecular interactions (bonded and nonbonded terms) are retained, this corresponds to exchanging a ligand from an ideal gas molecule state to a condensed phase and vice versa. This approach differs from absolute binding free energy schemes that use harmonic restraints or anchors to keep one ligand in place in the binding site[15,16,23] as the interactions of the fully interacting ligands at $\lambda = 0$ or $\lambda = 1$ with their surroundings are not biased by the present constraint. Where applicable, the simulations were also performed by a standard single-topology approach.

Results

Relative Free Energy of a Pair of Harmonic Oscillators. Before applying the above-described methodology in studies of biomolecular systems, it is important to verify that the simulation results are in agreement with other existing

methods. In particular, we wish to establish whether or not the calculated relative free energies need to be corrected because of the nature of the constraint employed in the dual-topology scheme described above. Because the free energies calculated in biomolecular systems can often exhibit a significant statistical uncertainty which could hide systematic differences, we first investigated a simple model of protein−ligand binding for which a relative binding free energy can be determined analytically.

Consider the following system:

$$A_1,P + A_2,0 \leftrightarrow A_1,0 + A_2,P \quad (4)$$

A_1 and A_2 are single atoms. They interact with environment P through a harmonic potential of identical equilibrium length $r_0 = 2.0$ length units and different force constants K_1 and K_2 of values 20 and 10 kcal/mol/length unit, respectively. Thus, A_1 and A_2 "bind" differently to P. The symbol $A_1,0$ means that A_1 is a gas-phase-like environment where it does not experience any intermolecular potential. When classical statistical mechanics is used and for this strictly one-dimensional problem, it is possible to derive an analytical expression for the relative free energy of A_2 and A_1.[42]

$$A_2 - A_1 = \Delta A = \frac{kT}{2} \ln\left(\frac{K_2}{K_1}\right) \quad (5)$$

At a temperature T of 298 K, this quantity amounts to −0.2053 kcal/mol. This relative free energy was estimated by three different free energy calculation techniques:

- *A single-topology scheme* where the force constant of the harmonic oscillator A_1 is changed by linear combination into the force constant of harmonic oscillator A_2.
- *The present dual-topology scheme* where both atoms are simulated simultaneously, but their interaction energy with P is scaled. In addition, both atoms are constrained to occupy the same position at every Monte Carlo move.
- *A double-decoupling approach* where the absolute binding free energies of A_1 and A_2 are first determined by application of the double-decoupling methodology.[21] The difference in their absolute binding free energies should give the correct relative binding free energy. Unlike the double-annihilation method,[43] in the double-decoupling approach, the decoupled ligand is restrained to occupy a well-defined region of space. The advantages of the restraint are two-fold. First, it makes the calculation reversible by avoiding the convergence issues associated with a ligand drifting out of a binding site. Second, because the volume the decoupled ligand occupies is defined by the restraint, the calculated absolute binding free energy can be related to a standard absolute binding free energy. Here, the ligands were restrained by a hardwall potential. When this form of restraint is used and according to Gilson et al.,[21] the absolute binding free energy can be calculated as

$$\Delta A_{\text{bind}}{}^{0} = \Delta A_{\text{bind}}^{\text{sim}} + kT \ln \frac{V_{\text{hardwall}}}{V_{\text{standard}}} \quad (6)$$

where $\Delta A_{\text{bind}}^{\text{sim}}$ is the free energy of turning off the interactions of the ligand with its environment in the presence of a hardwall and V_{hardwall} and V_{standard} are the hardwall and standard state volumes, respectively. For this simple, one-dimensional example, here, we arbitrarily define a standard "volume" of 2.0 length units. The simulation results are listed in Table 1. It can be seen that the single- and dual-topology schemes agree with the analytical solution. The backward runs are more precise than the forward run. This behavior was well interpreted by Kofke:[44] because A_1 has a larger force constant than A_2, the low-energy configurations of oscillator A_1 are a subset of the low-energy configurations of oscillator A_2, and sampling with the second oscillator gives a more thorough coverage of the configuration space.

Table 1. Free Energy of Two Harmonic Oscillators[a]

experiment	forward	backward
single topology	−0.2055 ± 0.0003	0.2052 ± 0.0001
dual topology	−0.2055 ± 0.0001	0.2054 ± 0.0001
decoupling A_1[b]	−1.8387 ± 0.0000	1.8387 ± 0.0000
decoupling A_2[b]	−1.8076 ± 0.0000	1.8076 ± 0.0000
ΔΔA[b]	0.0311 ± 0.0000	−0.0311 ± 0.0000
decoupling A_1[c]	−1.5255 ± 0.0028	1.5252 ± 0.0003
decoupling A_2[c]	−1.3217 ± 0.0002	1.3220 ± 0.0002
ΔΔA[c]	0.2038 ± 0.0028	−0.2032 ± 0.0003
decoupling A_1[d]	−1.3268 ± 0.0454	1.5248 ± 0.0008
decoupling A_2[d]	−1.2377 ± 0.0118	1.3197 ± 0.0010
ΔΔA[d]	0.0892 ± 0.0469	−0.2051 ± 0.0013
decoupling A_1[e]	−1.0204 ± 0.1285	1.5205 ± 0.0013
decoupling A_2[e]	−0.9004 ± 0.0484	1.3153 ± 0.0018
ΔΔA[e]	0.1200 ± 0.1373	−0.2052 ± 0.0022

[a] The figures are in kilocalories per mole and are the average of five independent single-window FEP simulations of 10M moves. One standard error is plotted as an estimate of the precision. Forward is for the perturbation of A_1 into A_2. Backward is for the perturbation of A_2 into A_1. For the absolute binding free energy calculations, a standard volume of 2.0 length units was arbitrarily defined. The hardwall was centered at the equilibrium bond length value of each oscillator. [b] A hardwall of width 0.1 length unit was applied. [c] A hardwall of width 0.5 length unit was applied. [d] A hardwall of width 1.0 length unit was applied. [e] A hardwall of width 2.0 length unit was applied.

In Table 1, the free energy changes for turning off one oscillator (decoupling A_1 or decoupling A_2) are also listed for different hardwalls. The difference between the free energy changes calculated for A_1 and A_2 and with the same hardwall should equal the relative binding free energy of both atoms. It can be seen that the hysteresis between the forward and backward free energy change is a function of the hardwall size. As expected, the backward runs are generally much more precise. From the absolute binding free energies obtained with the backward run, the relative free energy of A_1 and A_2 is not reproduced with the first hardwall, reproduced fairly with the second hardwall, and very well reproduced by the last two hardwalls. The behavior of the simulations is well understood in terms of the explanations put forward by Gilson and co-workers: the hardwall potential must not be too small so as to exclude the low-energy states of A1 or A2, but if it is too large, the simulations become less reversible and precise.[21]

This simple example demonstrates that the relative free energy calculated by a standard single-topology approach, the present dual-topology scheme, and a double-decoupling methodology (provided the hardwall is suitably chosen) all agree with the analytical result. There is therefore no

Table 2. Relative Free Energies Calculated for a Series of Molecular Systems[a]

experiment	solvent	coupling	soft core	free energy[b]
$\Delta\Delta G_{solv}$ ethane → methanol	TIP4P	single	NA	-5.95 ± 0.06[c]
$\Delta\Delta G_{solv}$ ethane → methanol	TIP4P	dual	$n = 0, \delta = 1.0$	-6.00 ± 0.05
$\Delta\Delta G_{solv}$ ethane → methanol	TIP4P	dual	$n = 1, \delta = 1.0$	-6.10 ± 0.06
$\Delta\Delta G_{solv}$ ethane → methanol	TIP4P	dual	$n = 2, \delta = 1.0$	-6.19 ± 0.10
$\Delta\Delta G_{solv}$ ethane → methanol	TIP4P	dual	$n = 1, \delta = 2.0$	-6.10 ± 0.16
$\Delta\Delta G_{solv}$ CE8 → CE1	TIP4P	single	NA	4.54 ± 0.04[c]
$\Delta\Delta G_{solv}$ CE8 → CE1	TIP4P	dual	$n = 0, \delta = 1.25$	4.61 ± 0.38
$\Delta\Delta G_{bind}$ CE8 → CE1	TIP4P	single	NA	-2.99 ± 0.07
$\Delta\Delta G_{bind}$ CE8 → CE1	TIP4P	dual	$n = 0, \delta = 1.25$	-2.72 ± 0.58
$\Delta\Delta G_{solv}$ CE8 → CE1	GBSA	single	NA	6.56 ± 0.01[c]
$\Delta\Delta G_{solv}$ CE8 → CE1	GBSA	dual	NA	6.53 ± 0.03
$\Delta\Delta G_{bind}$ CE8 → CE1	GBSA	single	NA	-2.52 ± 0.03
$\Delta\Delta G_{solv}$ CE8 → CE1	GBSA	dual	$n = 0, \delta = 1.25$	-2.82 ± 0.29
ΔG_{prot} CKA → CKB, inactivated	TIP4P	dual	$n = 0, \delta = 1.50$	7.97 ± 0.49
ΔG_{prot} CKA → CKB, inactivated	GBSA	dual	$n = 0, \delta = 1.50$	4.71 ± 0.29
ΔG_{prot} CKA → CKB, activated	TIP4P	dual	$n = 0, \delta = 1.50$	-5.80 ± 0.41
ΔG_{prot} CKA → CKB, activated	GBSA	dual	$n = 0, \delta = 1.50$	-4.56 ± 0.11
$\Delta\Delta G_{solv}$ CKB → BRY, activated	TIP4P	dual	$n = 0, \delta = 1.50$	3.07 ± 0.30
$\Delta\Delta G_{solv}$ CKB → BRY	GBSA	dual	NA	7.14 ± 0.01
$\Delta\Delta G_{bind}$ CKB → BRY	TIP4P	dual	$n = 0, \delta = 1.50$	-0.48 ± 0.50
$\Delta\Delta G_{bind}$ CKB → BRY	GBSA	dual	$n = 0, \delta = 1.50$	-5.62 ± 0.09

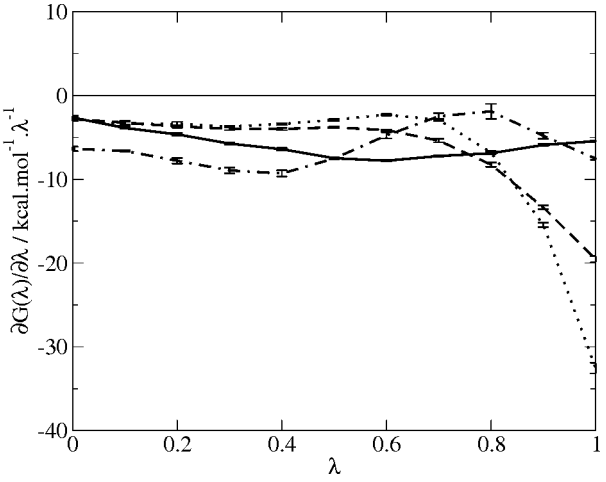
[a] The figures are in kilocalories per mole. $\Delta\Delta G_{solv}$ is a solvation free energy; $\Delta\Delta G_{bind}$ is a binding free energy, and ΔG_{prot} is the free energy difference between two ligands bound to a protein. [b] The error estimate is taken as one standard error from five independent simulations. [c] To obtain a relative solvation free energy with the single-topology method, the perturbation must also be carried out in vacuum. For the perturbation of ethane to methanol and CE8 to CE1, the free energy changes in the gas phase were respectively 2.69 ± 0.01 and 13.54 ± 0.01 kcal mol^{-1}

evidence from this system, or from the other systems that will be discussed later, that the present dual-topology scheme, in a Monte Carlo framework, requires a correction term arising from the applied constraints.

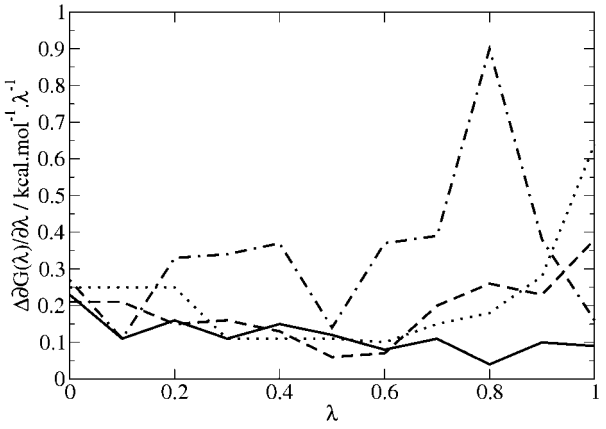
Relative Solvation Free Energy of Ethane and Methanol. The calculated relative solvation free energies of ethane and methanol for a series of single- and dual-topology calculations are listed in Table 2. It is apparent that, within statistical sampling error, the single- and dual-topology coupling schemes give the same free energy change. This suggests that the implementation of the dual-topology scheme is correct. It can also be seen that the parameters of the soft core influence the precision of the calculations. With δ set to 1.0, increasing n increases the spread of the individual simulation results. This can be understood by inspecting Figure 3a. The solute is perturbed from an apolar (ethane) to a polar (methanol) molecule, and as λ increases, it experiences stronger Coulombic interactions with the solvent. Increasing the parameter n results in the solute−solvent Coulombic interactions being restored later in the perturbation. This causes the free energy gradients to vary more rapidly in the second half of the perturbation. Because the free energy change is estimated by trapezium integration, smooth variations of the free energy gradients with λ should give more precise free energies, than profiles that change more rapidly. Also, as seen in Figure 3b and for this system at least, rapid variations of the free energy gradients are associated with larger error estimates and hence greater imprecision. When δ is increased from 1.0 to 2.0 with a constant value of n, the free energy gradients vary fairly smoothly, but the standard error on the estimate of the same free energy gradients obtained from five independent simulations fluctuates more, leading to increased imprecision. If δ is set to too high a value, then the solute−solvent Lennard-Jones energy is softened "too much" and the volume of space the solute can occupy varies more during the simulation, effectively making the perturbation more difficult. On the other hand, we have run a single simulation with δ set to 0.25 and n set to 1. The free energy change was -7.63 ± 0.57 kcal mol^{-1}, a figure that does not agree within error bounds with the previous free energy changes. Inspection of the free energy gradients reveals that most of the imprecision arises from widely fluctuating free energy gradients at the end of the perturbation. The difficulty here is that the soft core is not soft enough to allow a smooth decoupling of the molecule of ethane from the solvent, a problem typically observed in dual-topology simulations without a soft core.[12] Note that no such difficulty is observed at the beginning of the perturbation because the molecule of methanol fits inside the volume of ethane.

The present observations suggest that the soft-core parameters can be tuned to increase the precision of the calculated free energy change. In this process, analysis of the smoothness of the free energy gradient profiles and their statistical errors can be of valuable assistance.

Relative Binding Free Energy of Congeneric Inhibitors. The relative solvation free energy of two congeneric inhibitors of COX2[39,45] calculated by single- and dual-topology approaches is reported in Table 2. While the two methods give similar answers, it is clear that the single-topology calculations are much more precise. The same trend is observed for the relative binding free energy calculations: similar answers are obtained, but the single-topology calculations are 8−10 times more precise.



(a) Free energy gradients



(b) Error estimate on the free energy gradients

Figure 3. Ethane to methanol. (a) The free energy gradients for the perturbations carried out with various soft-core parameter sets. For the solid line, $n = 0$ and $\delta = 1.0$; for the dashed line, $n = 1$ and $\delta = 1.0$; for the dotted line, $n = 2$ and $\delta = 1.0$; for the dashed-dotted line, $n = 1$ and $\delta = 2.0$. (b) The standard error of the free energy gradients for the different parameter sets of the soft core.

The percentage of RETI moves that are accepted at each value of the coupling parameter λ is plotted in Figure 4. Because, in a RETI move, swaps of the configurations generated at neighboring values of λ are periodically attempted, the acceptance rate measures how much the equilibrium distributions of the neighboring replicas differ. As the free energy difference will converge more readily if the replicas are similar, a plot of this acceptance rate should give indications as to the difficulty in obtaining precise free energy differences along a given pathway. In the single-topology simulations, the replicas exchange more readily. In addition, the exchange rate is lower for the perturbation in the unbound state. This analysis corroborates the larger standard deviation of the dual-topology results. There remains the need to establish why the dual-topology simulations are much less precise. A plausible explanation can be put forward. In the single-topology simulation, the perturbation (mutation of a methyl group into a hydroxyl group) is well-localized on the scaffold. Other atoms in the molecule make a minor contribution to the free energy change (by small

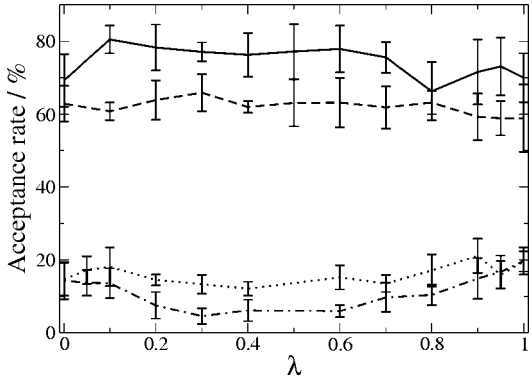


Figure 4. The acceptance rate of the RETI moves as a function of the coupling parameter λ for two congeneric inhibitors of COX2. The solid line is for the single-topology simulation of the bound state. The dashed line is for the single-topology simulation of the unbound state. The dotted line is for the dual-topology simulation of the bound state. The dashed-dotted line is for the dual-topology simulation of the unbound state. Each point is the average of five independent simulations, and the error bars show one standard deviation.

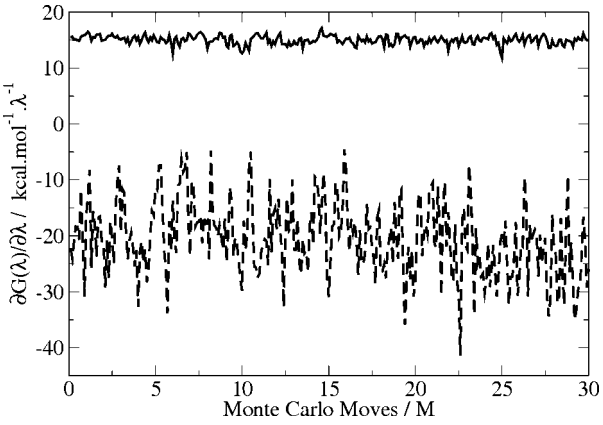


Figure 5. Fluctuations in the free energy gradients recorded during a Monte Carlo simulation. The data was extracted from an explicit solvent simulation perturbing the two congeneric inhibitors of COX2 bound to the protein at $\lambda = 0.50$. The solid line shows the gradients recorded with the single-topology protocol. The dashed line shows the gradient recorded with the dual-topology protocol.

variations of their atomic partial charges). In the dual-topology simulations, there are twice as many internal degrees of freedom to sample independently (those of each solute molecule) and the intermolecular interactions of every atom contribute directly to the free energy gradients. Thus, the dual-topology simulations should be intrinsically more difficult to carry out with sufficient precision. In Figure 5, the free energy gradients recorded by the single- or dual-topology technique at one value of the coupling parameter λ are plotted. It is evident that the free energy gradients recorded with the dual-topology technique fluctuate much more than those recorded with the single-topology technique. The data extracted from the simulation thus support the present explanation.

In Figure 5, it is apparent that the free energy gradients recorded with the single- and dual-topology techniques differ. This is because both techniques convert one ligand into

another by a different pathway, and in general, one does not expect similar gradients. The thermodynamic cycles shown in Figure 1 require only that the relative binding or solvation free energies be identical. Table 2 shows that, to within statistical error, this is indeed the case. Finally, in the previous system, ethane to methanol, no large difference in precision was observed between single- and dual-topology calculations. This is presumably because the numbers of atoms and internal degrees of freedom were much smaller.

Overall, these observations are in line with the findings of Pearlman that compared dual- and single-topology methods in a simple "ethane to ethane" perturbation[46] as well as comments from Shobana et al. who reported a hybrid single-/dual-topology technique.[27]

We have recently reported that binding free energy calculations can be carried out in an implicit solvent with good accuracy.[36] An implicit treatment of the solvent has the advantage of removing any sampling difficulties for the simulation in the unbound state, and the relative solvation free energies by the single- and dual-topology approaches are found once again to yield identical results to within a very narrow error interval (Table 2). There is still, however, a substantial difference in the precision of the relative binding free energy between the single- and dual-topology methods. In this system, the binding site is shielded from the solvent, and there is little precision to be gained for dual topology by adopting an implicit solvent model.

Binding Mode of a CDK2 Inhibitor in an Inactivated and Activated Complex. Cyclin-dependent kinase 2 (CDK2) plays an important role in the control of the cell cycle, is believed to be an important target for the development of cancer treatments, and is the focus of intense effort in drug development. CDK2 exists in an inactivated form, but the binding of cyclins A or E and subsequent phosphorylation of Thr160 causes significant conformational changes which greatly increase its phosphorylation activity.[47] Recently, Kontopidis et al. reported that a number of CDK2 inhibitors adopt different binding modes when complexed to an inactivated or activated CDK2.[41] Figure 6 shows the binding mode of 4-(2,4-dimethyl-1,3-thiazol-5-yl)pyrimidin-2-amine in inactivated and activated CDK2.

In inactive CDK2, the amino group of the ligand forms a hydrogen bond with the backbone carbonyl of Glu81. In the active complex, the amino group interacts with the backbone carbonyl of Leu83. This is achieved by a 180° flip of the pyrimidine ring. In addition, the thiazole ring moves substantially. Previous studies suggest that, in the inactivated form, favorable protein−ligand interactions are formed with Ile10 and Lys33.[40] In the activated form, the thiazole ring packs favorably against Phe80. The binding mode in inactivated and activated CDK2 will be referred to in the text as CKA and CKB, respectively. The flip of the pyrimidine ring, coupled with the translation and rotation of the thiazole ring, would be difficult to simulate with a single-topology method.

By contrast, setting up a dual-topology simulation is no more difficult than in the previous example. Here, we investigate with what accuracy the relative stability of each binding mode in both CDK2 forms can be predicted, and

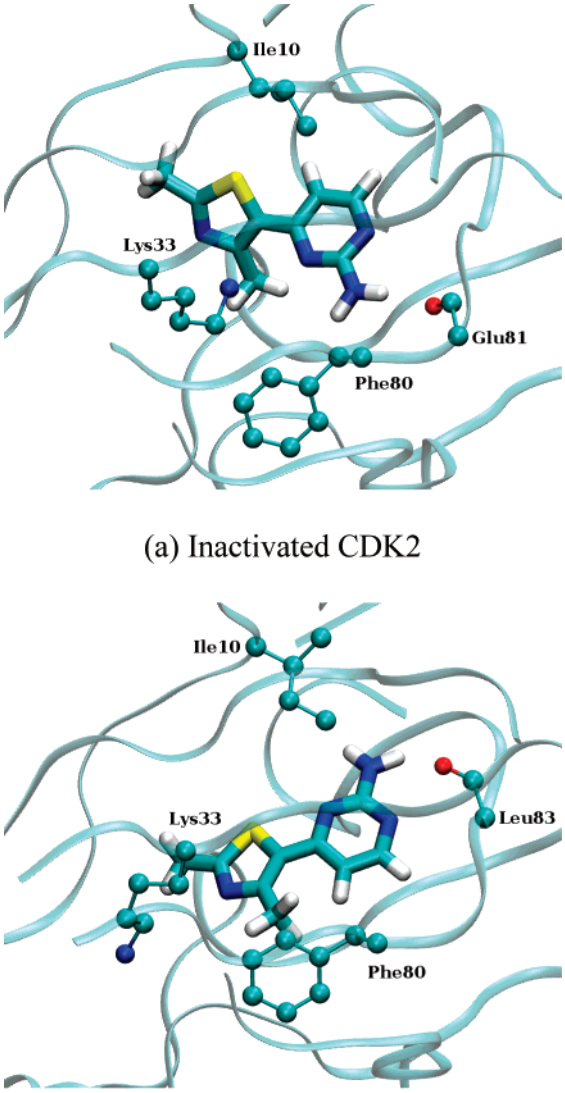


Figure 6. The binding mode of 4-(2,4-dimethyl-1,3-thiazol-5-yl)pyrimidin-2-amine in inactivated (top panel, PDB code 1PXJ) and activated (bottom panel, PDB code 2C5O) CDK2. The backbone carbonyl group of Glu81 and Leu83 is shown in CPK representation in the inactivated and activated CDK2, respectively. The side chains of Ile10, Lys33, and Phe80 are shown in CPK representation. The backbone of CDK2 is shown in ribbon representation. The ligand is shown in licorice representation. Hydrogen atoms on the protein are not shown, for clarity.

whether or not the results support the crystallographic evidence. The predicted relative free energies are listed in Table 2. The calculations were performed with explicit and implicit models of waters. Both types of calculation correctly identify the crystallographic binding mode as more stable, that is, that the CKA binding mode is favored in inactivated CDK2, while the CKB binding mode is favored in activated CDK2. The implicit solvent simulations are more precise than the explicit solvent simulations. This is because the water molecules in the vicinity of the ligand have to reorganize substantially to accommodate the change of binding mode. The difficulty of the explicit solvent simula-

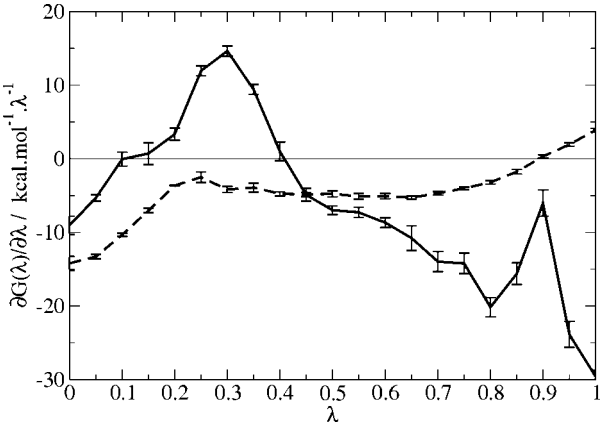


Figure 7. The free energy gradients recorded in the perturbation of CKA into CKB, bound to activated CDK2 (PDB code 2C5O). The solid line is for the explicit solvent simulation, and the dashed line is for the implicit solvent simulation. Error bars are shown for the free energy gradients obtained at each value of the coupling parameter λ.

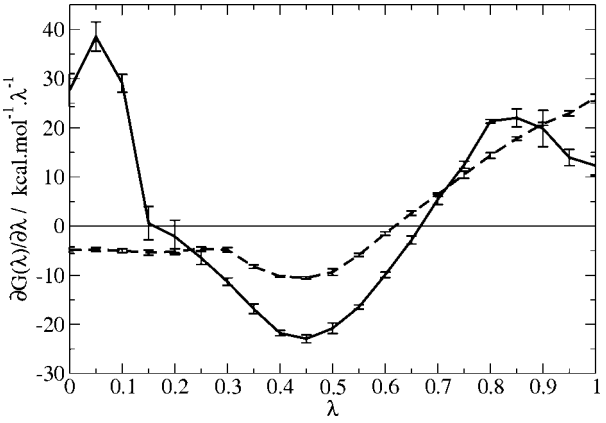


Figure 8. The free energy gradients recorded in the perturbation of CKB into BRY, bound to activated CDK2 (PDB code 2C5O). The solid line is for the explicit solvent simulation, and the dashed line is for the implicit solvent simulation. Error bars are shown for the free energy gradients obtained at each value of the coupling parameter λ.

tions is apparent when inspecting a plot of the free energy gradients in Figure 7. The implicit solvent simulations exhibit a smoother plot, with lower error bars on the individual free energy gradients. In addition, the acceptance rate of the RETI moves is on average twice as large as in the explicit solvent simulations (data not shown). However, the actual ΔG values differ, particularly in the inactivated CDK2. As no experimental measure of the relative stability of the two binding modes is available, it is not possible to conclude which protocol gives the more accurate answer. Certainly, one would not expect the behavior of the water molecules in the binding site to be well approximated by the GBSA model. However, the dual-topology method with both explicit and implicit solvation is able to assign the correct binding mode of the inhibitor to each kinase structure. This should be a simple test of the methodology as each protein structure is preorganized to stabilize a single binding mode. It is nevertheless important and comforting that the calculations are able to reproduce this trend clearly.

Relative Binding Free Energy of Two Different Scaffolds. The main objective of this research is to propose a method to calculate the relative binding free energy of structurally diverse molecules. To demonstrate this, we selected the inhibitor CKB described in the previous paragraph and attempted to perturb it into the CDK2 inhibitor 5-bromoindirubin (BRY) shown in Figure 2.[48] This system was selected because the two inhibitors share no common structural features and yet occupy the same position in the binding site. Relative solvation and binding free energies were calculated with implicit and explicit solvent models. The free energy gradients recorded for the perturbation in the bound state are shown in Figure 8. Observations similar to the previous system can be made. The free energy gradients in the explicit solvent simulations fluctuate more readily, and precise free energy estimates are harder to obtain than with the implicit solvent simulations. This is likely to be because transformation of CKB into the larger inhibitor BRY requires extensive solvent reorganization of the partially solvated binding pocket. Such a difficulty is of course not observed in the implicit solvent simulations. However, the binding free energies listed in Table 2 for each solvation model differ qualitatively. In the explicit solvent simulations, BRY and CKB are deemed inhibitors of similar potency to within the precision of the calculation, while BRY is predicted to be more potent by over 5 kcal mol^{-1} in the implicit solvent simulations. Such a large discrepancy might be related to the large difference in predicted relative solvation free energy. In implicit water, BRY is predicted 4 kcal mol^{-1} less stable than in explicit water; this would make the binding of BRY (which involves partial desolvation) more favorable. The reported IC50s for BRY and CKB are 1 μM and 6.5 μM, respectively, suggesting that BRY is a moderately more potent inhibitor.[48,49] However, the assay conditions differ, and not all the necessary parameters to permit conversion of the IC50s into inhibition constants were reported. Thus, rather than focusing on the accuracy of the results (which depends on the quality of the force field), we wish to emphasize the reproducibility of the calculations (which depends on the nature of the coupling scheme and the extent of sampling).

Discussion

In drug design, the identification of a promising scaffold ("hit") is often judged more difficult than the structure-based optimization of a given scaffold into a high-affinity compound ("lead"). Because of the computational expense required to obtain precise predictions and limitations in standard protocols, successful applications of free energy calculations have often been limited to lead optimization scenarios. The present methodology allows in principle for free energy differences between compounds of arbitrary shape to be calculated. This feature was highlighted by performing two different perturbations of CDK2 inhibitors that would have been difficult to set up in a single-topology paradigm. Because of ever increasing computational resources, sufficient precision in the predicted free energy changes to allow practical applications is now obtained at an affordable cost. Obvious applications include the inves-

tigation of the relative stability of different ligand binding modes predicted by docking programs, and significantly expanding the scope of free energy calculations in drug design by allowing consideration of structurally diverse compounds.

The setup of the perturbation is easier than in a single-topology approach, because it is no longer necessary to devise a scheme to convert gradually the topology of one ligand into the other. This makes it easier to set up rapidly a large number of compounds. The automation of such a task will be important to allow free energy calculations in a virtual screening context. However, before such an ambitious goal will be accomplished, a number of methodological challenges will have to be solved. The more the ligands differ in structure, the more likely it is that induced fit effects differ. Reliable free energy difference will only be obtained from a simulation if the sampling algorithms can relax the protein binding site sufficiently. In addition, structurally different ligands are more likely to have different low-energy conformers, when bound to the protein and in the aqueous phase. An accurate prediction of the free energy change will require thorough sampling of the energy minima of these conformers. This difficulty can be lessened by adopting an implicit solvent model, but with likely increased inaccuracy of the force field. Clearly, the force fields will have to accurately reproduce the energy difference between these minima to yield meaningful free energy changes. It has been suggested that this would be a major difficulty in absolute binding free energy calculations.[50] For flexible ligands, the complete sampling of their numerous energy minima will, in addition, require improved sampling algorithms. At the present, it is doubtful that Monte Carlo (or molecular dynamics) sampling methods can accomplish this generally. In these contexts, the proposed methodology could be used to test advanced sampling methods and calibrate force fields. Despite these limitations, a viable strategy for a free energy calculation drug design technology emerges from the results presented in this paper. When the present methodology is used, it is possible to calculate with reasonable precision the free energy change of small inhibitors of diverse structure. It could be therefore used to screen a number of low-molecular-weight scaffolds (e.g., heterocyclic rings). As these compounds are likely to have only a few rotatable bonds, existing sampling techniques are more likely to sample sufficiently the low-energy rotamers of the ligands in the unbound and bound states and yield precise, converged, results. In addition, if the screened compounds bind in the same part of the binding site, differences in induced fit effects should be lessened. Because this approach is compatible with popular developments in fragment-based screening technologies, it could be used to assist in the setup of fragment libraries. Once promising heterocycles are identified, substituent optimization should be accomplished by a single-topology method as more precise free energy estimates can be obtained for the same amount of computational resources. Alternative coupling schemes that combine dual- and single-topology features could also be envisaged, but such schemes would probably involve a complex system setup which would restrict their applicability.

Conclusion

We have described a methodology that allows the calculation of free energy differences between molecules of arbitrary shape and position. The methodology makes use of a dual-topology coupling scheme, a soft-core nonbonded energy function, and a constraint on the translation/rotation of the pair of solutes that can be easily implemented in a Monte Carlo simulation. Results identical to single-topology, double-decoupling, and analytical approaches are obtained for the relative free energy of a pair of harmonic oscillators. The method is as precise as a standard single-topology approach on the simple calculation of the relative solvation free energy of ethane and methanol. It proves less precise when applied to the calculation of the relative binding free energy of two inhibitors of COX2, although the two methods give identical results to within statistical error. However, as illustrated by two examples involving inhibitors of CDK2, the dual-topology method is readily applied to classes of problems that are beyond the reach of single-topology approaches. Precision can also be improved by adopting an implicit solvent approach, albeit at the expense of some accuracy. The computational expense is similar to standard single-topology approaches. This study highlights the strengths and weaknesses of single- and dual-topology methods for the calculation of relative free energies and suggests when one approach should be considered over the other. The present methodology demonstrates that relative binding free energy calculations between structurally diverse ligands can be computed with good precision and a reasonable amount of computational expense. As such, it should prove attractive to calculate relative binding free energies between sets of ligands that would have been previously only considered feasible by more challenging absolute binding free energy calculation methodologies. It is hoped that the present method will extend the scope of free energy simulations and find useful applications in drug design studies.

Acknowledgment. We thank the University of Southampton for funding this work and EPSRC (GR/R06137/01) for providing computational resources.

References

(1) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.

(2) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050−3054.

(3) Miyamoto, S.; Kollman, P. A. *Proteins: Struct. Funct. Bioinf.* **1993**, *16*, 226−245.

(4) Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663−9669.

(5) Fox, T.; Scanlan, T. S.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 11571−11577.

(6) Price, M. L. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 9455−9466.

(7) Udier-Blagovic, M.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2004**, *47*, 2389−2392.

(8) Jorgensen, W. L. *Science* **2004**, *303*, 1813−1818.

(9) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069−1072.

(10) Simonson, T. Chapter 9. In *Computational Biochemistry and Biophysics*, 1st ed.; Marcel Dekker: New York, 2001; p 169.

(11) Beutler, T. C.; Mark, A. E.; Vanschaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.

(12) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025−9031.

(13) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 103−118.

(14) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 119−136.

(15) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535−9551.

(16) Roux, B.; Nina, M.; Pomès, R.; Smith, J. C. *Biophys. J.* **1996**, *71*, 670−681.

(17) Schafer, H.; van Gunsteren, W. F.; Mark, A. E. *J. Comput. Chem.* **1999**, *20*, 1604−1617.

(18) Pitera, J. W.; van Gunsteren, W. F. *J. Phys. Chem. B* **2001**, *105*, 11264−11274.

(19) Oostenbrink, C.; van Gunsteren, W. F. *J. Comput. Chem.* **2003**, *24*, 1730−1739.

(20) Oostenbrink, C.; van Gunsteren, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6750−6754.

(21) Gilson, M. K.; Given, J. A.; Bush, B.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.

(22) Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825−6830.

(23) Deng, Y. Q.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255−1273.

(24) Sleigh, S. H.; Tame, J. R. H.; Dodson, E. J.; Wilkinson, A. J. *Biochemistry* **1997**, *36*, 9747−9758.

(25) Barillari, C.; Taylor, J.; Viner, R.; Essex, J. W. *J. Am. Chem. Soc.* **2007**, *129*, 2577−2587.

(26) Hornak, V.; Okur, A.; Rizzo, R.; Simmerling, C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 915−920.

(27) Shobana, S.; Roux, B.; Andersen, O. S. *J. Phys. Chem. B* **2000**, *104*, 5179−5190.

(28) Woods, C. J.; Michel, J. *ProtoMS2.1*; in-house Monte Carlo software, 2006.

(29) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703−13710.

(30) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(32) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129−152.

(33) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.

(34) Michel, J.; Taylor, R. D.; Essex, J. W. *J. Comput. Chem.* **2004**, *25*, 1760−1770.

(35) Michel, J.; Taylor, R. D.; Essex, J. W. *J. Chem. Theory Comput.* **2006**, *2*, 732−739.

(36) Michel, J.; Verdonk, M. L.; Essex, J. W. *J. Med. Chem.* **2006**, *49*, 7427−7439.

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

(38) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132−146.

(39) Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; McDonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Miyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C. *Nature* **1996**, *384*, 644−648.

(40) Wang, S. D.; Meades, C.; Wood, G.; Osnowski, A.; Anderson, S.; Yuill, R.; Thomas, M.; Mezna, M.; Jackson, W.; Midgley, C.; Griffiths, G.; Fleming, I.; Green, S.; McNae, I.; Wu, S. Y.; McInnes, C.; Zheleva, D.; Walkinshaw, M. D.; Fischer, P. M. *J. Med. Chem.* **2004**, *47*, 1662−1675.

(41) Kontopidis, G.; McInnes, C.; Pandalaneni, S.; McNae, L.; Gibson, D.; Mezna, M.; Thomas, M.; Wood, G.; Wang, S.; Walkinshaw, M.; Fischer, P. *Chem. Biol.* **2006**, *13*, 201−211.

(42) McQuarrie, D. A. *Statistical Mechanics*, 1st ed.; Harper and Row: New York, 1976.

(43) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742−3746.

(44) Kofke, D. A. *Mol. Phys.* **2004**, *102*, 405−420.

(45) Marnett, L. J.; Kalgutkar, A. S. *Curr. Opin. Chem. Biol.* **1998**, *2*, 482−490.

(46) Pearlman, D. A. *J. Phys. Chem.* **1994**, *98*, 1487−1493.

(47) Collins, I.; Garrett, M. D. *Curr. Opin. Pharmacol.* **2005**, *5*, 366−373.

(48) Jautelat, R.; Brumby, T.; Schafer, M.; Briem, H.; Eisenbrand, G.; Schwahn, S.; Kruger, M.; Lucking, U.; Prien, O.; Siemeister, G. *ChemBioChem* **2005**, *6*, 531−540.

(49) Wu, S. Y.; McNae, I.; Kontopidis, G.; McClue, S. J.; McInnes, C.; Stewart, K. J.; Wang, S. D.; Zheleva, D. I.; Marriage, H.; Lane, D. P.; Taylor, P.; Fischer, P. M.; Walkinshaw, M. D. *Structure* **2003**, *11*, 399−410.

(50) Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2006**, *49*, 5880−5884.

CT700081T