Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics

Guha Jayachandran
Computer Science Department, Stanford University, Stanford, California 94305

Michael R. Shirts
Chemistry Department, Columbia University, New York, New York 10027

Sanghyun Park and Vijay S. Pande[a)]
Chemistry Department, Stanford University, Stanford, California 94305

(Received 6 April 2006; accepted 16 June 2006; published online 22 August 2006)

We present a technique for biomolecular free energy calculations that exploits highly parallelized sampling to significantly reduce the time to results. The technique combines free energies for multiple, nonoverlapping configurational macrostates and is naturally suited to distributed computing. We describe a methodology that uses this technique with docking, molecular dynamics, and free energy perturbation to compute absolute free energies of binding quickly compared to previous methods. The method does not require *a priori* knowledge of the binding pose as long as the docking technique used can generate reasonable binding modes. We demonstrate the method on the protein FKBP12 and eight of its inhibitors. © *2006 American Institute of Physics*.
[DOI: 10.1063/1.2221680]

I. INTRODUCTION

The calculation of free energy changes associated with molecular processes has long been a goal of computational chemistry.[1] Free energy perturbation (FEP) and its close relative, thermodynamic integration (TI), are prominent physics-based methods of computing absolute free energies of processes such as small molecule solvation or protein-ligand binding.[2,3] While much work has gone into improving their efficiency, there are still significant hurdles preventing the routine evaluation of the free energy of most biological interactions. These hurdles can be divided into two categories: weaknesses in the model (force field, solvent treatment, etc.) and shortcomings in sampling. The methods presented here are aimed only at improving the latter.

Sampling biomolecular systems is a challenge as the probabilistically relevant regions of conformational space may be large or may include large barriers. With molecular dynamics simulation, the time scale to cover all probabilistically relevant conformations with an equilibrium distribution may be on the microsecond or millisecond time scale, while molecular dynamics simulations of proteins generally only reach the nanosecond time scale.[4] With Monte Carlo simulation, the major alternative to molecular dynamics, an efficient move set for branched biopolymers, particularly one that can handle explicit water correlated motions, is not apparent.

The sampling problem is exacerbated when we lack structural information about the equilibrium state. Even if the most probabilistically relevant positions comprise only a small portion of the total configurational space (for example, for binding, a highly restricted binding mode), these positions may not be known if no cocrystal or other experimental structural information is available. This common scenario requires sampling over a large number of configurations to ensure that the unknown binding mode is captured.

Short of developing new hardware or simulation models that can sample more rapidly, a natural line of attack is parallelization. The weighted histogram analysis method (WHAM) with umbrella sampling is one established method that can be parallelized, for example.[5] Parallelized algorithms are especially attractive, given the increased adoption of grid and distributed computing in recent years. Here, the method presented offers benefits even if not parallelized, but is ideally suited for highly parallel architectures and naturally scales to thousands of processors.

Our proposed method has three aims. The first is to dramatically shorten wall clock time for sampling by using only short (50 ps) individual trajectories (even at the cost of increasing the number of trajectories). The second aim is to reduce the importance of *a priori* knowledge of the experimental end states. In the context of ligand binding, this means making experimental knowledge of the binding mode unnecessary to perform free energy calculations. The final aim of the method is operational flexibility and ease. In particular, it should not require manual selection of any coordinate or require overlap of predefined states, unlike WHAM or other restraint methods. It furthermore should operate easily with explicit solvent, without the necessity of designing a move set to handle the correlated solvent motions.

In implementing the method, we must generate an initial diverse set of configurations, ideally including ones in high equilibrium probability regions such as binding modes. As we elaborate below, we apply docking to this problem. Docking has been used for two decades in drug discovery efforts.[6] It has two main uses: prediction of bound protein-

[a)]Electronic mail: pande@stanford.edu

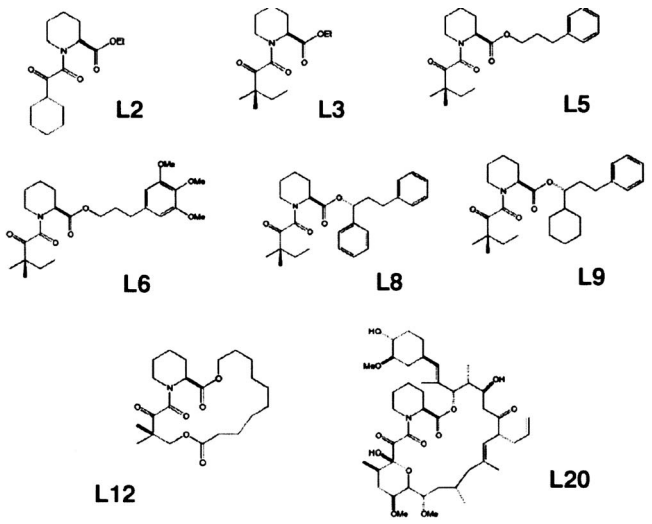

FIG. 1. The eight ligands whose binding with FKBP was considered. Their free energies of binding were all experimentally measured by Holt *et al.*, and the naming follows that work. L20 is FK506, a naturally occurring inhibitor of FKBP that is utilized pharmaceutically.

ligand conformations and scoring of binding affinities, both across different poses of a single ligand and across different ligands. Here, we use docking only for generating complex conformations, with molecular mechanics, free energy perturbation, and the Bennett acceptance ratio[7,8] (BAR) used to rigorously calculate the binding affinity.

We demonstrate the parallelized technique developed by computing the binding free energies of eight ligands with FKBP12 (Fig. 1). While assessing FKBP binding in detail is not a focus of this work, a reasonable agreement with experimental measurements was achieved. FKBP and its ligands have been used as a model system in several previous computational studies,[3,9,10] as the protein is relatively small in size (107 residues), involves relatively little conformational change, and offers a set of ligands of varying sizes and affinities that have been experimentally characterized.[11]

II. THEORY

In free energy perturbation, to obtain the free energy difference between two states A and B, we break the process of transforming from the A Hamiltonian to the B Hamiltonian into discrete stages called lambda intervals. Lambda (λ) is a parameter that determines the degree to which the Hamiltonian has transformed ($\lambda=0$ is pure state A and $\lambda=1$ is pure state B). We can denote the change in free energy over a given interval by $\Delta G_{\lambda_i\rightarrow\lambda_{i+1}}$. Given forward and backward potential energy differences between the end states of an interval, the maximum likelihood $\Delta G_{\lambda_i\rightarrow\lambda_{i+1}}$ over the interval can be determined using BAR,[7] otherwise, thermodynamic integration or exponential averaging of work measurements in a single direction may be used. The derivation presented here is independent of which approach is used, but we use BAR due to its greater efficiency.[7,8,12] The desired overall free energy change is simply the sum over the intervals:

$$\Delta G=\sum_{i=0}^{N-1}\Delta G_{\lambda_i\rightarrow\lambda_{i+1}}, \tag{1}$$

where $\lambda_0=0$ and $\lambda_N=1$. The reason for the use of lambda intervals is to increase efficiency and precision (by increasing the overlap between ensembles at neighboring lambda values), but it also means that sampling for FEP can be naturally parallelized over lambda intervals.

The technique described here further parallelizes over configurational states. It is similar in spirit to previous works which sampled different conformational regions separately.[13] We begin the derivation by introducing the notation we shall use for describing free energy changes. $\Delta G(\lambda_i,X\rightarrow\lambda_j,Y)$ will denote the change in free energy associated with a move from configurational state X in the λ_i Hamiltonian to configurational state Y in the λ_j Hamiltonian. We define a configurational state to be a set of microstates meeting some configurational conditions—thus X and Y are each a set of microstates. As we will describe in Sec. III, the definition of configurational states can be performed based on the obtained simulation data rather than *a priori*. The free energy associated with the entire transformation described in Eq. (1) is denoted $\Delta G(0,U\rightarrow 1,U)$, where U is the set of all microstates of the system. Expanding this ΔG into a summation over the lambda intervals as before, we obtain

$$\Delta G(0,U\rightarrow 1,U)=\sum_{i=0}^{N-1}\Delta G(\lambda_i,U\rightarrow\lambda_{i+1},U). \tag{2}$$

Now, we rewrite $\Delta G(\lambda_i,U\rightarrow\lambda_{i+1},U)$ in terms of partition functions.

$$\Delta G(\lambda_i,U\rightarrow\lambda_{i+1},U)=-kT\ln\frac{\sum_{\mu\in U}w_{\lambda_{i+1}}(\mu)}{\sum_{\mu\in U}w_{\lambda_i}(\mu)}, \tag{3}$$

where k is the Boltzmann constant, T is the temperature, and $w_\lambda(\mu)$ is the Boltzmann weight of microstate μ in Hamiltonian λ.

Working on just the fraction from the right side of Eq. (3), we partition the numerator:

$$\frac{\sum_{\mu\in U}w_{\lambda_{i+1}}(\mu)}{\sum_{\mu\in U}w_{\lambda_i}(\mu)}=\frac{\sum_{s\in R}\sum_{\mu\in s}w_{\lambda_{i+1}}(\mu)}{\sum_{\mu\in U}w_{\lambda_i}(\mu)}=\sum_{s\in R}\frac{\sum_{\mu\in s}w_{\lambda_{i+1}}(\mu)}{\sum_{\mu\in U}w_{\lambda_i}(\mu)}. \tag{4}$$

R is a set of mutually exclusive subsets of U that collectively cover U exhaustively; s is one of those subsets. Now, we will derive an expression for the quotient on the right side of Eq. (4) that is amenable to direct computation. First, we rewrite it in terms of a free energy change:

$$\frac{\sum_{\mu\in s}w_{\lambda_{i+1}}(\mu)}{\sum_{\mu\in U}w_{\lambda_i}(\mu)}=\exp(-\Delta G(\lambda_i,U\rightarrow\lambda_{i+1},s)/kT). \tag{5}$$

Next, we break $\Delta G(\lambda_i,U\rightarrow\lambda_{i+1},s)$ into three pieces:

$$\Delta G(\lambda_i, U \to \lambda_{i+1}, s) = \Delta G(\lambda_i, U \to 0, U) + \Delta G(0, U \to 0, s) + \Delta G(0, s \to \lambda_{i+1}, s) \tag{6}$$

$$= -\Delta G(0, U \to \lambda_i, U) + \Delta G(0, U \to 0, s) + \Delta G(0, s \to \lambda_{i+1}, s).$$

We will address the three terms in the above expression in reverse order. The third term is simply

$$\Delta G(0, s \to \lambda_{i+1}, s) = \sum_{j=0}^{i} \Delta G(\lambda_j, s \to \lambda_{j+1}, s), \tag{7}$$

where each term in the sum is simply an output from BAR or an exponential averaging on the given lambda interval.

The second term in Eq. (6) is not as straightforward, as it involves the change from all of configurational space U to a restricted subset s of microstates. In a very simple case where the configurational state is determined entirely by a molecule's translational position rather than any internal degrees of freedom and the Hamiltonian is center-of-mass translationally invariant, only entropy needs to be considered:

$$\Delta G(0, U \to 0, s) = -kT \ln(V(s)/V(U)), \tag{8}$$

where $V(X)$ is the volume of conformational state X. $V(s)$ is determined by the definition of s. $V(U)$ depends on how the overall end states are defined. Later, we describe one approach (using a Markov model) that will work regardless of how states are defined and will allow states defined based on internal degrees of freedom. The key fact aiding us will be that the transformation at $\lambda = 0$ can be defined to make the computation of a restriction free energy at that Hamiltonian particularly tractable.

Returning to Eq. (6), the first term is the negative of

$$\Delta G(0, U \to \lambda_i, U) = \sum_{j=0}^{i-1} \Delta G(\lambda_j, U \to \lambda_{j+1}, U), \tag{9}$$

where each term of the sum (which is over lambda intervals preceding λ_i) has been computed using Eq. (3). Note that evaluation of Eq. (9) through a given lambda interval requires results of Eq. (3) only for preceding lambda intervals. Thus, if we evaluate the summation in Eq. (2) in sequential order, we can easily obtain $\Delta G(0, U \to 1, U)$.

The above derivation shows that if one partitions configuration space into regions ($s \in R$), the free energy change can be independently computed for each region and the results then assembled together. In this approach, a weighted average over configurational states is taken for each Hamiltonian step, and these averages are then summed to obtain the free energy change associated with the entire process. A similar alternate that naturally comes to mind is to reverse the order—to first get a sum for each configurational state over all Hamiltonian steps (in λ) and then take a weighted average of those overall sums. In particular, we start with

$$\Delta G(0, U \to 1, U) = -kT \ln \frac{\sum_{\mu \in U} w_1(\mu)}{\sum_{\mu \in U} w_0(\mu)}.$$

Breaking the numerator into subsets as before gives

$$-kT \ln \sum_{s \in R} \frac{\sum_{\mu \in s} w_1(\mu)}{\sum_{\mu \in U} w_0(\mu)} = -kT \ln \sum_{s \in R} \exp(-\Delta G(0, U \to 1, s)/kT).$$

Using a thermodynamic cycle, we obtain

$$-kT \ln \sum_{s \in R} \exp[(-\Delta G(0, U \to 0, s) - \Delta G(0, s \to 1, s))/kT],$$

which finally yields

$$\Delta G(0, U \to 1, U) = -kT \ln \sum_{s \in R} \left[\exp\left(-\Delta G(0, U \to 0, s) - \sum_{i=0}^{N} \Delta G(\lambda_i, s \to \lambda_{i+1}, s)/kT \right) \right]. \tag{10}$$

In fact, this equation can be shown to be equivalent to the formulation provided earlier [Eqs. (2)–(6)].

In standard FEP, one must sample the overall equilibrium distribution at each intermediate λ state. In the formulation described above, we only need to sample the overall equilibrium distribution at one Hamiltonian ($\lambda = 0$ in the derivations above). At all the other λ values, we need only sample within each configurational state. This allows for trivial parallelization over the configurational states in sampling. Furthermore, given that we do not need to observe any transitions between states (except at $\lambda = 0$), the reduction in wall clock time for obtaining sufficient sampling can be expected to be greater than linear in the number of partitions for a highly parallel implementation of the method.

III. METHODS

In this section, we describe a methodology for utilizing the above parallelized-over-parts FEP (POPFEP) to calculate free energies of binding. The absolute free energy of binding for a ligand can be obtained by summing the free energy changes associated with two processes that together constitute binding: desolvation of the ligand from bulk solvent to an ideal gas and solvation of the ideal gas ligand into the context of the protein binding pocket.[3,14] Each of those two quantities is computed independently and, as will be seen below, the overall scheme is the same for both situations. The methodology, in either case, can be divided into five

steps: (1) generation of initial configurations, (2) FEP molecular dynamics simulations, (3) configurational state assignment, (4) computation of state probabilities at $\lambda=0$, and (5) combination of per state free energies into an overall result.

The goal of step 1, generation of initial configurations, is to obtain a diverse seeding of conformational space. The second step, FEP simulations, provides potential energy differences between neighboring Hamiltonians for many sampled conformations. Step 3, configurational state assignment, assigns sampled conformations into configurational states s which, together with the output of the preceding step, provides all necessary data for calculation of $\Delta G(\lambda_i, s \rightarrow \lambda_{i+1}, s)$ values with BAR. Step 4, computation of state probabilities at $\lambda=0$, provides $\Delta G(0, U \rightarrow 0, s)$ values. All the values required for evaluation of Eq. (10) are thus procured, and all that remains in step 5 is to obtain a value for $\Delta G(0, U \rightarrow 1, U)$.

Below, we will describe a practical implementation of the methodology and application to the example of FKBP12 and eight of its known ligands (Fig. 1). The same ligands were considered as in the Fujitani *et al.* study.[10] The same initial protein structures were also used. In particular, for each ligand of interest, the protein structure of FKBP from the cocrystal with a ligand most similar to the ligand of interest was used. Protein Data Bank (PDB) structure 1FKF (FKBP bound FK506) was used for L20 simulations. Protein from 1FKG (FKBP bound L8) was used for L8. Protein from 1FKH (FKBP bound L9) was used for L2, L3, L5, and L6. Protein from 1FKI (FKBP bound L13) was used for L12.

A. Generation of initial configurations

Input: Protein structures and ligands.

Output: Diverse set of ligand conformations, in complex and in solution.

The first step in the method is to generate initial configurations. Choosing configurations such that low free energy regions are represented will dramatically reduce the sampling required, as we will not need to wait for molecular dynamics to reach them from other regions. In the context of binding, such key configurations are binding poses.

1. Configurations of complexes

We docked ligands to FKBP crystals, specified above, to obtain initial binding poses. Docking was performed using Surflex v.1.23.[15] No information on the binding pocket was inputted for the PROTOMOL generation step, but the program properly identified the pocket in all cases. Thorough search parameters were used and the top eight scoring poses were retained. We carried the top eight scoring poses from docking to the FEP stage to use as initial structures for simulation and as seeds for generation of further initial structures.

As noted, by choosing diverse initial configurations, we have reason to believe we can sample space with shorter simulation trajectories than otherwise. To supplement the configurations obtained from docking, therefore, we also generated some random bound poses. Each of the docked complexes was "boiled" to generate such additional poses. Eight trajectories starting from each complex were run at 500 K for 700 ps with the protein frozen. A simple distance-dependent dielectric implicit solvent was used (similar to Ref. 16). Given the goal of randomization, the quality of the solvent model was relatively unimportant as long as it did not impede sampling. Speed therefore motivated the choice. Structures were sampled at a 100 ps frequency. The 512 structures sampled per ligand (8 docked results $\times 8$ trajectories $\times 8$ samples per trajectory) were pruned to discard overly similar conformations: We retained conformations generated such that all retained had a root mean square deviation (RMSD) of at least 1.5 Å with any other retained complex or any complex from docking. The RMSD considered was in ligand coordinates with the fit of two complexes based on protein coordinates so as to capture translational differences, relative to the pocket, between two ligand poses.

For FKBP, the docking and the high temperature pose generation resulted in a total of 9–51 initial configurations, generally varying with ligand size. While protein conformational flexibility was not addressed here, we note that the same approach could be extended to generate variety in protein conformations, not just in ligand conformations. Such an extension would be most practical for proteins where the flexible portion is limited, such that protein structure can be included as part of the configurational state definitions (discussed in Sec. III C).

2. Configurations of ligands

To compute absolute binding free energies, we need to compute ligand solvation energies and thus need ligand-alone initial configurations as well. For these configurations, we utilized the identical ligand conformations obtained above, simply removing the protein. This choice may not be ideal, given that the distribution of ligand microstates is different depending on whether it is in the context of bulk solvent or protein. An alternative approach, both for generating ligand-alone conformations and for generating bound poses, would be to use a software (such as OpenEye's OMEGA[17]) that generates low energy, druglike ligand conformations. Parallel tempering techniques may also be suitable. Another, more comprehensive, alternative would be to systematically vary torsional angles of the ligand and generate a regularly spaced coverage of conformational space—however, this could result in unmanageable numbers of poses depending on the spacing.

B. FEP simulations

Input: Initial configurations for complex and ligand.

Output: Potential energy differences between neighboring Hamiltonians for many sampled conformations.

From each of the initial conformations generated in the preceding step, we conducted molecular dynamics FEP simulations as described in Ref. 10. Individual trajectories for the complex were of length 100 ps, and for the ligand alone were of length 1 ns. Data from 30 to 50 ps were used for primary analysis and to full length in examining convergence (Sec. IV). These are lengths achievable in less than a day on a standard contemporary CPU, making the computation

theoretically possible in a day with massive parallelization to handle the tens of starting configurations used for each ligand.

1. Molecular dynamics details

The same force field parameters were used as in the previous work.[10] In particular, AMBER 99-Φ (Ref. 18) was used for the protein and ions. The general AMBER force field[19] (GAFF) provided ligand parameters except for charges, which came from MOPAC 2002 using AM1-BCC.[20] TIP3P explicit solvent was used.[21] The simulations were run on the Folding@Home distributed computing infrastructure.[22]

Simulations were run using a modified form of GROMACS 3.1.4,[23] in double precision, adapted to the Folding@Home environment. The velocity Verlet algorithm was used for integration with a 2 fs time step.[24] All bonds were constrained using SHAKE/RATTLE with a relative tolerance of 10^{-12} (square of bond length deviation). The Andersen thermostat (with an all-atom velocity reassignment every 2 ps) was utilized to maintain a temperature of 298 K.[25] The Parrinello-Rahman barostat was used with pressure set to 1.01 bar.[26] A neighbor list of 10 Å was utilized, with an update frequency of 20 fs. van der Waals (vdw) interactions were switched between 8 and 9 Å distances. Particle mesh Ewald[27] (PME) with an interpolation order of 4 was used for long range Coulombic interactions, beyond 9 Å. The Fourier spacing was approximately 1.2 Å.

All simulations were run in a truncated octahedron box. The complex boxes were sized for the minimum distance between an atom and an image atom to be at least 1.2 Å. For the ligand boxes, that value was 1.4 Å. There were 4000–5000 water molecules in the complex simulations and 300–800 water molecules in the ligand-only simulations. The complex simulations included four chlorine ions to neutralize the charge from the protein.

2. Free energy perturbation

Our FEP involved sampling at each of a number of lambda values and measuring the potential energy difference from transforming to adjacent lambda values (both higher and lower).[7,8] Potential energy differences between Hamiltonians were evaluated every 0.1 ps. We simulated four trajectories from each starting configuration at each lambda. As discussed earlier, lambda values in our case correspond to turning off interaction between the ligand and its surroundings.

As in Refs. 3 and 10, one set of Hamiltonians used corresponded to linear variation of the Coulombic interactions between the ligand and surroundings, and another set of Hamiltonians used corresponded to varying degrees of the vdw interaction between the ligand and surroundings using a soft-core potential with Coulombic interactions fully off. The lambda values used were as follows, where 0 corresponds to no interaction and 1 to full interaction: 0, 0.35, 0.55, 0.73, 0.88, and 1 for the Coulombic set and 0, 0.1, 0.2, 0.25, 0.3, 0.4, 0.55, 0.7, 0.85, and 1 for the vdw set. A reproduction of the FKBP computation[10] of Fujitani *et al.* showed these values to yield similar errors as those used in that work.

3. Complex decoupling end state

In computing ligand solvation free energies, the two end states are well characterized—ligand fully solvated and ligand as an ideal gas, interacting with itself as normal but not with solvent. For the other process whose free energy is computed—desolvation of the ligand from the context of the binding pocket—one end state is ligand fully interacting with protein and solvent. To reach the other end state, however, there are two possible approaches: decoupling of the ligand from the solvent and protein through an intermediate restraint and decoupling from solvent and protein without a restraint.[28]

When restraints are used, the ligand is typically attached to the protein binding pocket with a spring. Traditionally termed "double decoupling," this method has the advantages of requiring sampling of the "ligand as an ideal gas" end state only within the range of the spring and of having a clear connection with standard state conditions.[28] Decoupling the ligand without a restraint does not utilize a spring. This has traditionally been called "double annihilation" (as in Gilson *et al.*[28] and Fujitani *et al.*[10]), but we will avoid the use of that term as the word "annihilation" has also been used to mean the loss of ligand-ligand interactions in addition to ligand-environment interactions.

Decoupling with a restraint can require less sampling for convergence. However, here we have followed the work of Fujitani *et al.* with FKBP (Ref. 10) and used decoupling without a restraint. Simulation without restraint fits more naturally into our overall methodology, as conformations can be assigned to states without regard to the point to which they were restrained and without any steps required to compute free energies of restraint, as described in the next section.

A weakness of decoupling without a restraint can be the lack of connection with standard state conditions. In this work, we address this through the use of an alternate thermodynamic cycle. This cycle, detailed in Sec. III F, formally requires the decoupled end state of the complex decoupling step to be an ideal gas ligand still within the binding site. This differs from the work of Fujitani *et al.* in that the volume of the ideal gas is clearly defined.

The binding site volumes were decided as follows. After collecting the simulation data, we computed for all sampled conformations the distance between center of mass of the ligand and the binding pocket. The anchor point in the binding pocket was defined for FKBP as the center of mass of the alpha carbons of SER38, VAL55, and TRP59. We listed the distances observed at the fully coupled Hamiltonian in increasing order and took the value at the 95% percentile as defining the cutoff for the bound state. At loosely coupled Hamiltonians, some conformations were outside the threshold and discarded as not being part of the bound end state. The volume of the state was the volume of a sphere with a radius equal the threshold. A correction connecting this volume with standard state conditions is given in Sec. III F.

C. State assignment

Input: Sampled conformations (associated with potential energy differences).

Output: Configurational state assignments for sampled conformations.

From the simulations above, we have a number of conformations and potential energy differences for these conformations to neighboring lambda values. Our next task is to partition the conformations into the configurational states that were introduced earlier. It is informative to consider the extreme cases of state partitioning. If there is just one large state, then the method recovers standard FEP. If, on the other hand, each conformation is its own state, then we face the problem that we will only have the free energy differences associated with two lambda intervals (up one and down one from the Hamiltonian the conformation was generated under) rather than for all the lambda intervals comprising the full process.

1. Characterization and dimensionality reduction

In this work, we partitioned conformations into states as follows. First, for each ligand conformation, we computed pairwise distances between all nonhydrogen atoms. For the case of the complex, we also included in the computation the three atoms from the protein binding pocket specified in the preceding section, to capture the translational position of the ligand relative to the pocket. If n was the number of atoms considered, then a vector of $n(n-1)/2$ values was obtained and used to characterize the conformation.

Next, we pruned the dimensionality of our representation. This has two benefits. One is vastly increasing computational efficiency versus using the full dimensionality representation for each conformation, which can be intractable in some cases. For example, by reducing the dimensionality to four for the steps below, we made linear operations on the vectors approximately 100 times faster than on the full dimensionality representations (which ranged in size from 190 to 1770). Note that if too few dimensions are retained, then distinguishability between conformations could be lost. The other benefit of dimensionality reduction is focusing the characterization of states to kinetically relevant dimensions, such that members of a given state can interconvert easily compared to members of different states. We chose the dimensions based only on data from the Hamiltonian where the ligand was fully interacting with its surroundings ($\lambda=1$), as that is where sampling was presumably slowest.

We wished to retain those dimensions where motion is slow but where a wide range of motion is possible. The first step in the selection process was to calculate, for each trajectory that was obtained under the fully coupled Hamiltonian, the mean and standard deviation for each vector element over all vectors from that trajectory. We then averaged the standard deviations over all the fully coupled trajectories. This value (s) is a rough measure of the kinetic stability of a given interatomic distance, with lower standard deviation meaning more stable. We also computed for each vector coordinate the largest difference between its maximum trajectory average and minimum trajectory average. This value (r) gives a sense of the overall range available to the coordinate. We computed the ratio r/s for each dimension and ranked them from highest to lowest, so that those dimensions ranked highest where the observed spread was a high multiple of the per-trajectory standard deviation. Finally, we stepped through the ranking and discarded entries corresponding to an interatomic distance involving an atom that was already involved in a higher ranking dimension. The restriction of allowing each atom to be involved in only one retained dimension aims to reduce correlation between retained dimensions.

2. Clustering

Given reduced dimension representations of each conformation, our next step was to use those representations for state construction. To do this, we first clustered a subset of the vectors corresponding to the fully coupled Hamiltonian data to obtain seed states. We took data at a 5 ps resolution from each trajectory run under the fully coupled Hamiltonian from 30 to 50 ps for a total of $20N$ conformations, where N is the number of starting conformations for the given ligand (5 conformations/ trajectory $\times$ 4 trajectories/starting conformation $\times N$ starting conformations). These were clustered with hierarchical clustering.[29] The average linkage criteria with a 0.5 Å cutoff was used.

Next, the remaining conformations, from all Hamiltonians, at the full time resolution, were assigned to the nearest seed cluster, with the distance between a conformation and a cluster deemed to be the Cartesian distance between its vector representation and the mean vector (center) of the cluster. Finally, we checked if any clusters lacked a member from any of the lambda intervals. If so, we merged it with the cluster whose center was nearest to its own, repeating until all clusters had representation from all lambda intervals. This procedure resulted in between 3 and 10 configurational states for ligands alone and between 23 and 35 states for complexes.

3. Alternatives

There are many alternate approaches possible for assigning conformations to states. We have experimented, for example, with the use of supervised machine learning techniques like support vector machines.[29] The optimal way of defining states should be explored further and, indeed, the current active study of Markovian state decomposition for protein dynamics could provide useful techniques here.[30,31]

We divide approaches to partitioning into two categories: descriptive and prescriptive. The approach outlined above is an example of what we call a descriptive approach, with the states defined based on the data observed. In what we call a prescriptive approach, the states are defined before any simulation is conducted (see Ref. 32 for a current work exemplifying this approach). Restraints may be used to force given trajectories to sample only within assigned predefined states. A key advantage of this approach over a descriptive one is that it will be easier to judge whether each state has reached equilibrium.

However, unless we have the computational power to systematically define and sample a comprehensive set of states, it may be difficult to decide *a priori* how best to define the states for molecules with many degrees of freedom, and clearly, the introduction of poorly constructed restraints could prevent sampling of key phase space regions. Restraints creating flat bottom wells can be easily and simply constructed, and would likely lessen this risk compared to harmonic restraints, but they would also not offer as much benefit as more sophisticated restraints in terms of focusing the sampling. Indeed, flat bottom wells would in practice make little change to our current scheme, given the variety of starting positions. With some types of constraints, measuring state probabilities at the fully decoupled Hamiltonian (essentially measuring the free energy of restraint) could require much sampling. Further work should be done to compare descriptive and prescriptive approaches to see whether either has an edge in practical applications and to see how they compare over a range of systems. Hybrid approaches—for example, where states are prescriptively defined before sampling and descriptively redefined after sampling—are also possible.

D. State probabilities at $\lambda=0$

Input: Protein structure and initial ligand conformations.

Output: Probability of each defined configurational state at $\lambda=0$ Hamiltonian.

As noted in Sec. II, the POPFEP formalism does not require that we obtain equilibrium distributions over the entire configurational space at all Hamiltonians, but does require it for at least one. In practice, this is easiest to do at the fully decoupled Hamiltonian ($\lambda=0$), as that is where sampling is most rapid. In particular, in our binding example, the ligand in the fully decoupled state can be modeled as a ligand in vacuum. Even for the complex case, we only require data from the ligand's dynamics in vacuum as it does not interact with protein or solvent, and thus simulation is very fast. We further discuss this point below, explaining how we can account for the position of the free ligand relative to the protein. We used a Markov model approach to obtain the state probabilities but replica exchange, stochastic roadmap simulation, or other techniques may well prove more efficient.

Markov models have been recently used to study protein dynamics.[31,33,34] They are defined by a transition matrix where element (i,j) of the matrix gives the probability of the system being in state j some step time after being seen in state i. They do not require equilibrium sampling over the entire space but only sufficient sampling between neighboring states to obtain accurate transition probabilities. Here, the transition probabilities between configurational states over a step time of 10 ps were computed based on the transitions observed in the simulations described below. The first eigenvector (associated with the eigenvalue 1) of the transition matrix gave the equilibrium probabilities of the states. The full procedure of constructing a matrix is described in both Refs. 31 and 34.

To obtain greater sampling than was obtained from the FEP simulations in the decoupled Hamiltonian, we simulated four 10 ns trajectories of a ligand in vacuum from each of the previously determined starting conformations (the first 1 ns was discarded). For computing ligand solvation energy, sampled conformations (at a frequency of 10 ps) were assigned to configurational states as part of the procedure described in Sec. III C. For computing complex decoupling energy, the procedure is described in the following paragraph.

As mentioned earlier, even for the protein-ligand complex case, it should be sufficient to sample the ligand on its own, as it does not interact with the rest of the system. The only issue, however, is that our characterization of complex configuration includes the ligand's position relative to the pocket (as noted in Sec. III C). To obtain this relative position from simulations of ligand alone, we take the trajectories sampled above and break them into 1 ns blocks. We call the initial conformation of the trajectory S. At every 1 ns interval of the trajectory, we fit the ligand conformation sampled at that point to the ligand conformation in the initial bound complex from which S was taken. We retain the protein coordinates from the fit and treat them as fixed relative to the ligand for the next 100 ps of the trajectory under consideration. In this way, we can capture motion of the ligand relative protein and take into account the translational dimensions of states. The remaining 900 ps from every 1 ns block is not used. This keeps the data to a manageable amount. Also, at a longer time after a fit, the ligand will have drifted away from the protein and will no longer be in the binding pocket. We could have used snapshots from 1 ns continuously instead of spread over 10 ns, but the latter offers increased coverage of phase space; the additional simulation is acceptable given its speed (few hours on typical computers).

E. Combination to overall free energy

Input: Many sampled potential energy differences between Hamiltonians and state assignments for sampled conformations.

Output: Free energy of full transformation, $\Delta G(0,U \rightarrow 1,U)$.

Having completed the preceding steps, we could now compute the overall free energy for each process (desolvation and complex decoupling). Potential energy differences were used in BAR at their sampled 0.1 ps frequency while, for computational reasons, we had assigned conformations to states at only a 1 ps resolution. We associated the ten work values from the picosecond following a sampled conformation with the state of that conformation (using five values preceding and following a sampled conformation works similarly). With BAR, we obtained a free energy for each lambda interval for each state. We could then evaluate Eq. (10) with these free energies to obtain the overall free energy of the process. In this way, for each ligand we obtained a free energy for desolvation and a free energy for coupling the ligand with protein and solvent.

Given the configurational states defined, statistical error estimates were obtained through a bootstrap analysis. For each lambda interval and configurational state pair, we com-

puted a new ΔG value with BAR using a random selection of the total pool of work measurements available at that lambda interval and state. The number of work measurements selected was equal to the number of work measurements available, with the random selection of blocks of ten consecutive measurements done with replacement. We note that in a bootstrapping calculation, the block size should be twice the correlation time of the values. If there are some states with a very long correlation time, the error could be underestimated. The difficulty in assessing correlation in each state is a weakness of the method used. The ΔG values calculated with the randomly selected potential energy difference measurements for each state and lambda interval were combined as described earlier [Eq. (10)] to obtain a total ΔG for the process, either ligand solvation or complex decoupling. This procedure was repeated 20 times and the standard deviation over the trials was computed. We note that this is a statistical error only—it does not account for model errors. Also, if the sampling was poor in the sense that it left important regions entirely unexplored, that fact would not be reflected in this error.

F. Sum for binding free energy

Input: ΔG of ligand solvation, ΔG of decoupling ligand from complex, and binding site volume.

Output: ΔG for binding.

Having obtained free energies for its constituent processes of ligand solvation and decoupling from complex, we used them to obtain the free energy of binding. In traditional double decoupling or double annihilation,[28] there is no explicit definition of a binding volume required. A Boltzmann distribution is formally obtained for the entire space. A restraining spring ameliorates the issue of having to sample the entire simulation box by making sampling of the range of the spring formally sufficient, and the presence of this spring is accounted for analytically and a connection with standard state volume and concentration established. In the present work, however, a definition of the binding pocket is required (described in Sec. III B) in order to make a connection to the standard state.

In particular, we correct for the entropy difference between the binding site volume and the standard state volume with $-kT\ln(V^b/V^o)$, where V^b is the binding site volume and $V^o=1660$ Å^3, the volume for one molecule in $1M$ solution. In practice for our simulations, the correction ranged from -0.04 to -0.64 kcal/mol. Summarizing then, the free energy of binding was considered to be the sum of the free energies of following thermodynamic steps (L refers to ligand and P to protein).

(1) $L(\text{coupled}) \rightarrow L(\text{decoupled})$ [ligand desolvation energy; formally considered to be in V^o and distant from protein].
(2) $L(\text{decoupled, in } V^o) \rightarrow L(\text{decoupled, in } V^b)$ [accounted for by $-kT\ln(V^b/V^o)$].
(3) $P+L(\text{decoupled, in } V^b) \rightarrow P+L(\text{coupled, in } V^b)$ [ligand coupling energy in the context of protein].

These sum to the overall process

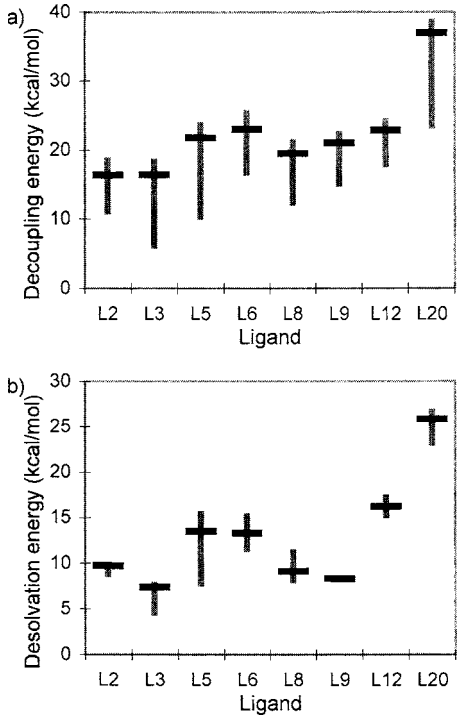

FIG. 2. The computed free energies of (a) complex decoupling and (b) desolvation for each ligand. The range of the minimum to maximum free energy observed for any configurational state, $\Delta G(0,s\rightarrow 1,s)$, is shown. The horizontal bars denote the final POPFEP outputs.

$$P + L(\text{coupled, far from protein}) \rightarrow P + L(\text{coupled, in binding pocket}).$$

The variance in the total binding energy is taken to be the sum of the variances in the ligand solvation energy and complex coupling energy.

IV. RESULTS

We carried out the above method for FKBP and eight of its ligands. This set includes both manmade and naturally occurring ligands. Holt *et al.* designed some of those ligands and measured the inhibition constant K_i for all.[11] Fujitani *et al.* approximated the dissociation constant K_d to equal K_i (the approximation is most applicable for strongly binding ligands) and converted it to a free energy using

$$\Delta G_{\text{exp}} = kT\ln K_d. \tag{11}$$

The assays of Holt *et al.* were at 283 K, but $T=298$ above to match the temperature of the simulations, the temperature for which the force field is most appropriate.[10] The validity of this approximation is unclear, but it appears the most reasonably possible for the available data.

A. Free energies of binding

Figure 2 shows the range of desolvation and decoupling energies obtained for configurational states for each ligand, along with the result of combining with Eq. (10). Evaluating the sum of the calculated desolvation, coupling, and volume correction terms, as described in the preceding section, yielded overall binding free energies. Figure 3 shows a graph

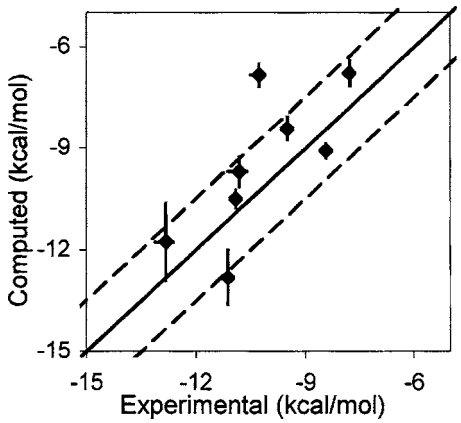

FIG. 3. Computed binding free energies using POPFEP vs experimentally measured binding free energies for eight FKBP ligands. The lines $y=x$ and $y=x\pm1.5$ are drawn as guides. The most outlying point is associated with L12.

comparing the experimental binding free energies with those computed in this work (also in Table I). The RMSD between the two sets is 1.6 kcal/mol. Ligand L12 is the major outlier, its computed free energy deviating from the experimental value by a substantial 3.4 kcal/mol. A better choice of starting configurations, as discussed in Sec. III A, may improve the situation. The RMSD excluding L12 is 1.1 kcal/mol.

We stress that, though the results appear predictive, we make no claim that the agreement will generalize to other systems. We point out that very minor changes in the model—for example, in the ligand atom charges—can cause changes of a kcal/mol or more in solvation free energy or decoupling energy.[35] Our results here, like those in most computational studies utilizing molecular mechanics force fields, may well be benefiting from a virtuous cancellation of error between the solvation free energy term and the complex decoupling term.

The statistical error in computed values (Table I), described earlier, is in the tenths of a kcal/mol and is largest for the largest ligands. In Fig. 2, the almost invisible error bars corresponding to experimental errors are simply the error reported by Holt *et al.*. They do not reflect the systematic error that is manifested in the fact that different experimental techniques or groups can obtain slightly different free energies for a single process. Similarly, the error bars for the computation do not reflect systematic errors from the model, which are likely of higher magnitude.

TABLE I. The binding free energies obtained from experimental values and the computed free energies. All values are in kcal/mol.

	Computed			
Ligand	Solvation	Complex coupling plus volume correction	Binding	Experimental binding
L2	−9.74±0.06	−16.5±0.37	−6.78±0.37	−7.8
L3	−7.43±0.05	−16.5±0.21	−9.08±0.21	−8.4
L5	−13.4±0.09	−21.8±0.34	−8.42±0.35	−9.5
L6	−13.3±0.18	−23.0±0.41	−9.71±0.45	−10.8
L8	−9.13±0.07	−19.6±0.25	−10.5±0.26	−10.9
L9	−8.31±0.12	−21.1±0.80	−12.8±0.81	−11.1
L12	−16.2±0.14	−23.0±0.31	−6.84±0.34	−10.3
L20	−25.8±0.17	−37.6±1.13	−11.8±1.14	−12.8

TABLE II. The ligand solvation, complex coupling, and overall binding free energies reported in Ref. 10. The RMSDs of the solvation, complex coupling, and binding free energy column with the corresponding columns in Table I are 1.9, 4.0, and 2.8 kcal/mol, respectively. We emphasize that direct comparison of complex coupling or binding energies is complicated by the fact that the protocol of Ref. 10 did not make connection with the standard state volume.

Ligand	Solvation	Complex coupling	Binding
L2	−8.5	−12.8	−4.3
L3	−7.6	−12.5	−4.9
L5	−8.4	−14.7	−6.3
L6	−12.7	−20.9	−8.2
L8	−9.3	−16.6	−7.3
L9	−7.1	−15.3	−8.2
L12	−15.4	−22.5	−7.1
L20	−26.0	−36.1	−10.1

B. Single configuration decoupling

It is desirable to compare in more detail starting simulation from a single configuration versus starting from many configurations. We focus the comparison on complex decoupling as that is where equilibration is slower and thus where differences in methodology would be expected to be more pronounced. In the earlier work by Fujitani *et al.*, a single initial complex structure, obtained from tens of nanoseconds of molecular dynamics equilibration (prior to any free energy perturbation), was utilized for each ligand's free energy simulation (Table II).[10] (If a constant offset of 3.2 kcal/mol in the favorable direction was added to the computed binding free energies, then the RMSD of this fit with experiment was 0.4 kcal/mol; this was without any volume correction or restraint, and the connection with the standard state is unclear.) If a single initial configuration has to be chosen, lengthy equilibration is a fairly rigorous method of selecting it. It is therefore worthwhile to include the initial complexes from that previous work in analysis of the effect of starting configuration, particularly as one motivation of the present work was to avoid such equilibration. We will sometimes refer to the set of initial conformations from the work of Fujitani *et al.* as simply the "equilibrated" structures or conformations.

To compare the equilibrated structures with data from the current work, we first have to account for the fact that the previous work utilized a different molecular dynamics protocol. Notably, it utilized a different integrator, different pressure control algorithm, and different constraints algorithm, though the same force field parameters.[10] We repeated the complex decoupling calculations of the previous work, following its single starting configuration protocol and using the same initial conformations, but using the molecular dynamics model of the present work. The results obtained, using sampling from 600 to 1000 ps of each of ten trajectories for each Hamiltonian, had a RMSD of 0.8 kcal/mol with the decoupling energies from the previous work.

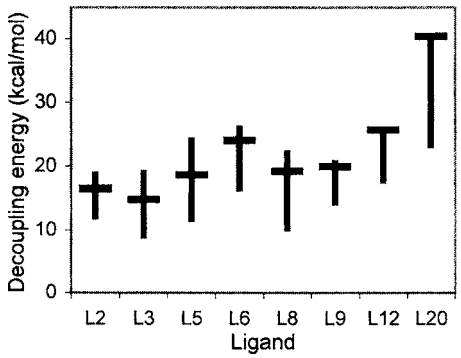

FIG. 4. Treating each set of simulations we had begun from a given conformation independently, the minimum to maximum decoupling energies observed. The horizontal bars denote the decoupling free energies obtained from using the starting conformations of Ref. 10.

TABLE III. Complex decoupling energies obtained using only simulations from an initial complex obtained through lengthy structural equilibration (structures from Ref. 10) and the decoupling energies obtained using Eq. (10) with simulations began at multiple structures. (All using data between the 30 and 50 ps point of trajectories.) All values are in kcal/mol. the RMSD between the two columns is 2.1 kcal/mol.

Ligand	Single equilibrated complex	POPFEP
L2	16.5	16.3±0.37
L3	14.8	16.4±0.21
L5	18.6	21.7±0.34
L6	24.1	22.9±0.41
L8	19.2	19.5±0.25
L9	19.9	21.0±0.80
L12	25.7	22.8±0.31
L20	40.4	36.9±1.13

We were now better able to undertake a comparison between data from the equilibrated conformations and other single initial conformation results. For each of the starting configurations used in our POPFEP computation, we used the already collected data to calculate a decoupling energy by the simple single-state protocol of the earlier work, so that one decoupling energy was independently obtained from each starting configuration's data. We used only data from 30 to 50 ps from each trajectory and did the same for the equilibrated initial structures' trajectories, making for data sets that can be obtained in equal wall clock time (ignoring preparation of structures). We also did not define a restricted binding volume. The methods followed for obtaining all the decoupling energies (one from the earlier work's single initial conformation and one from each of the present work's initial conformations) thus were identical except for the starting conformation used. Note that the trajectory length does not allow as much drift away from the binding pocket as the longer trajectories of the work of Fujitani *et al.* (where no restraints were used or correction for volume made),[10] contributing to values of greater magnitude than those seen in that work. This difference does not concern us here as our goal in the remainder of this section is not to compare to that work's energies but to examine the impacts of initial structure on the shorter time scale used in this study.

Figure 4 shows the decoupling free energy obtained from the equilibrated single conformation in the context of the minimum and maximum decoupling free energies obtained from the current work's initial conformations. One might expect that the earlier work's initial conformations would yield decoupling energies of near or greater magnitude as the maximum decoupling energies from the present work's configurations, as the earlier conformations were obtained by a long equilibration period that presumably would have culminated in a location with favorable energy. This is indeed what we see for the largest ligands, but for the smaller five ligands, and especially for L3 and L5, the decoupling energies from the earlier work's structures lie in the middle of the ranges. One possibility is that smaller ligands may be allowed more motion in the pocket, and a long preparatory trajectory may terminate in an entropically favorable but energetically nonoptimal location, causing free energy simulation on the order of a nanosecond or less from that position to neglect conformations that are of lower potential energy. We point out that this is not necessarily harmful to accuracy in practice, as it may be the case that such neglected regions are in actuality of low probability: indeed, certain ligands yield very similar decoupling energies to the POPFEP result even with only 500 ps of sampling (10 trajectories $\times$50 ps) per Hamiltonian (Table III). However, the fact that the range between minimum and maximum observed single-state decoupling energy is relatively large for each ligand (Fig. 4) highlights the fact that if one is simulating from a single starting structure, great care (as in the form of a lengthy structural equilibration) must be taken in selecting that structure and lengthy simulation may be necessary beyond that in order to fully sample phase space with proper probabilities.[3,13]

Finally, we note further results that agree with the fact that it is crucial to consider entropy, not just potential energy. In particular, we examine the effect of considering from our overall POPFEP calculation only the single defined configurational state (as per Sec. III C) for each process—solvation and complex decoupling—that saw the largest free energy change $\Delta G(0,s\rightarrow 1,s)$ (Fig. 2). The resulting binding free energy (including same volume corrections as before) shows poorer agreement with the experimental binding free energies than the result utilizing all states. The RMSD with experiment is 2.0 kcal/mol with affinities generally overly favorable compared to experiment, suggesting that taking into account multiple configurational states (motion) is worthwhile for FKBP and that only sampling the very local vicinity of locations with the most favorable free energy is insufficient.

C. Longer simulation

Conducting very long simulations and comparing the structures sampled there with those sampled in the present work would be highly informative as to the extent and importance of conformational space coverage. Such long trajectories are of course highly difficult to obtain, however, especially for complexes, one of the motivations for the method shown here. While structural analysis is beyond the scope of the present work, we are currently pursuing such an experiment with collaborators.

TABLE IV. Ligand solvation energies and complex coupling energies (plus volume correction) obtained from longer simulation. Data from 600 to 1000 ps of each trajectory was used for the ligands and from 30 to 100 ps for the complexes. All values are in kcal/mol. The same data is presented in Figs. 5 and 6 in comparison to the results using less data.

Ligand	Solvation	Complex coupling plus volume correction
L2	−9.19±0.02	−16.1±0.26
L3	−6.94±0.01	−15.8±0.12
L5	−10.4±0.02	−20.3±2.12
L6	−12.7±0.01	−22.4±0.22
L8	−9.21±0.04	−18.9±0.14
L9	−7.94±0.02	−19.3±0.13
L12	−16.2±0.10	−23.0±0.15
L20	−25.6±0.06	−36.7±0.33

Longer trajectories can also provide information on convergence within the model. In the above discussed results, we only used data from between the 30 and 50 ps time points of each trajectory. However, our ligand-alone simulations ran to 1 ns. We recomputed ligand solvation energies using the same method as described earlier (including redoing state definition) except that we used only data from between the 600 ps and 1 ns time points from FEP trajectories (Table IV). Figure 5 shows the solvation free energies computed using the 30–50 ps data versus those computed using the 600–1000 ps data. The RMSD between the two sets is 1.1 kcal/mol. The outlier is ligand L5, showing a difference of nearly 3 kcal/mol; excluding L5 results in a RMSD of 0.4 kcal/mol. The divergence in L5 appears due to differences in the states defined, but it is yet unclear what about L5's structure causes this. Though far slower than ligand-only simulation, we were also able to extend our complex trajectories to 100 ps to get an idea of how those energies change (Table IV). Figure 6 shows the coupling free energies (including standard state volume correction) computed using the 30–50 ps data versus those computed using the 30–100 ps data. The RMSD between the two sets is 1.0 kcal/mol, with all the energies except one agreeing within one standard deviation error. We stress that, given the length of simulation, these results do not at all prove convergence of the free energies—they are necessary but insufficient tests.

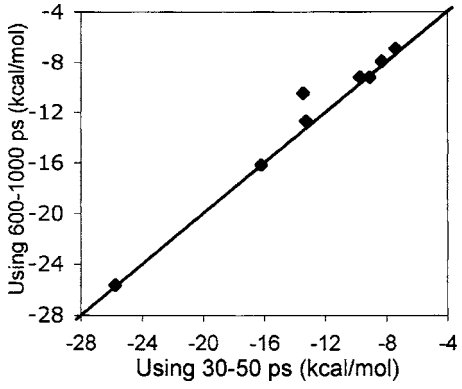

FIG. 5. Solvation free energies computed using data from between time points of 30 and 50 ps of the FEP trajectories and using data between time points of 600 ps and 1 ns (error bars too small to be seen). The line $y=x$ is drawn as a guide. The outlier is L5.

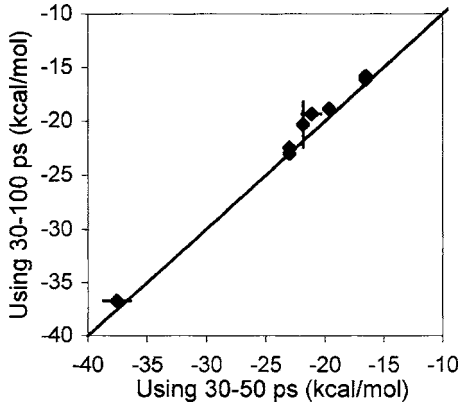

FIG. 6. Free energy of decoupling the ligand within the binding pocket (including volume correction), as computed using data between time points 30 and 50 ps of trajectories or using data between 30 and 100 ps of trajectories. All ligands show agreement within error except for L9.

V. CONCLUSION

We have presented a formalism for computing free energies over disjoint configurational states with FEP and combining them. It is very naturally adapted to parallel computing (including loosely coupled CPUs as found in distributed computing), offering the promise of reducing wall clock times, which has been a key obstacle to the wider scale use of FEP. It also reduces the importance of *a priori* structural knowledge of the end states.

In the context of binding, we described a method using docking to help generate initial configurations for simulation. This use of docking aids sampling of favorable conformations and, together with the configurational state averaging technique described, reduces dependence on knowledge of the true end states. We also described an automated approach for assigning sampled conformations to configurational states based on significant dimension representations. Other approaches for this step can be used without alteration of the rest of the method.

Applying the described method to computing absolute binding free energies for FKBP resulted in reasonable agreement with experiment, although, as discussed, it is difficult to make any judgment about agreement with experiment in other systems without further study. The average sampling used for each ligand, including both solvation and complex decoupling computations, was 200 ns, underscoring the difficulty of applying the method without parallelization. The computed statistical error was under 1 kcal/mol for most ligands. Longer simulation or more trajectories may help improve the precision and further test convergence. Better methods of identifying slow degrees of freedom and using them to define states may also help.

Our hope is that methods such as that described can increase the practicality of free energy computations. Not only is reducing wall clock time important for its own sake, but it also makes possible computations that previously were not. Of course, studying larger systems is one example of this, but we can also contemplate the use of more detailed (and typically considerably slower) simulation models such as a polarizable force field or hybrid quantum mechanics and molecular mechanics (QM/MM) methods. With sampling

less of an issue, models can be improved faster. Together, improved sampling and modeling hold the promise of dramatically improving the utility of free energy computation in drug discovery.

ACKNOWLEDGMENTS

The authors thank Folding@Home participants around the world. They thank Hideaki Fujitani for providing the ligand parametrizations used in this study and also John Chodera and David Mobley for their discussions. One of the authors (M.R.S) is supported by a NIH NRSA fellowship. This work was supported by a grant from NSF for cyberinfrastructure.

[1] P. A. Kollman, Chem. Rev. (Washington, D.C.) **7**, 2395 (1993); C. Chipot and D. A. Pearlman, Mol. Simul. **28**, 1 (2002).
[2] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[3] M. R. Shirts, Ph.D. dissertation, Stanford University, 2005 (ProQuest location: http://wwwlib.umi.com/dissertations/fullcit/3153076).
[4] T. Hansson, C. Oostenbrink, and W. F. v. Gunsteren, Curr. Opin. Struct. Biol. **12**, 190 (2002); M. Karplus and J. A. McCammon, Nat. Struct. Biol. **9**, 646 (2002).
[5] M. Souaille and B. Roux, Comput. Phys. Commun. **135**, 40 (2001); S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comput. Chem. **13**, 1011 (1992); C. Bartels, M. Schaefer, and M. Karplus, J. Chem. Phys. **111**, 8048 (1999).
[6] B. D. Bursulaya, M. Totrov, R. Abagyan, and C. L. Brooks III, J. Comput.-Aided Mol. Des. **17**, 755 (2003); G. L. Warren, C. W. Andrews, A.-M. Capelli *et al.*, J. Med. Chem. (2005).
[7] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976); M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, Phys. Rev. Lett. **91**, 140601 (2003).
[8] M. R. Shirts and V. S. Pande, J. Chem. Phys. **122**, 144107 (2005).
[9] M. L. Lamb and W. L. Jorgensen, J. Med. Chem. **41**, 3928 (1998); M. L. Lamb, J. Tirado-Rives, and W. L. Jorgensen, Bioorg Med. Chem. **7**, 851 (1999).
[10] H. Fujitani, Y. Tanida, M. Ito, G. Jayachandran, C. D. Snow, M. R. Shirts, E. J. Sorin, and V. S. Pande, J. Chem. Phys. **123**, 084108 (2005).
[11] D. A. Holt, J. I. Luengo, D. S. Yamashita *et al.*, J. Am. Chem. Soc. **115**, 9925 (1993).
[12] M. R. Shirts, J. W. Pitera, W. C. Swope, and V. S. Pande, J. Chem. Phys. **119**, 5740 (2003).
[13] A. Hodel, L. M. Rice, T. Simonson, R. O. Fox, and A. T. Brunger, Protein Sci. **4**, 636 (1995); M. Leitgeb, C. Schroder, and S. Boresch, J. Chem. Phys. **122**, 84109 (2005); T. P. Straatsma and J. A. McCammon, *ibid.* **90**, 3300 (1988).
[14] S. Boresch, F. Tettinger, M. Leitgeb, and M. Karplus, J. Phys. Chem. A **107**, 9535 (2003).
[15] A. N. Jain, J. Med. Chem. **46**, 499 (2003).
[16] M. Y. Shen and K. F. Freed, Proteins **49**, 439 (2002).
[17] OpenEye scientific software, OMEGA, http://www.eyesopen.com/products/applications/omega.html
[18] E. J. Sorin and V. S. Pande, J. Comput. Chem. **26**, 682 (2005); W. D. Cornell, P. Cieplak, C. I. Barly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[19] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25**, 1157 (2004).
[20] A. Jakalian, D. B. Jack, and C. I. Bayly, J. Comput. Chem. **23**, 1623 (2002).
[21] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[22] M. Shirts and V. S. Pande, Science **290**, 1903 (2000); http://folding.stanford.edu
[23] E. Lindahl, B. Hess, and D. v. d. Spoel, J. Mol. Model. **7**, 306 (2001).
[24] M. Amini, J. W. Eastwood, and R. W. Hockney, Comput. Phys. Commun. **44**, 83 (1987).
[25] H. C. Andersen, J. Chem. Phys. **52**, 24 (1980).
[26] S. Nose and M. L. Klein, Mol. Phys. **50**, 1055 (1983); M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).
[27] T. Darden, D. York, and L. Pederson, J. Chem. Phys. **98**, 10089 (1993).
[28] M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, Biophys. J. **72**, 1047 (1997).
[29] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 2002); T. M. Mitchell, *Machine Learning* (McGraw-Hill Higher Education, New York, 1997).
[30] W. C. Swope, J. W. Pitera, and F. Suits, J. Phys. Chem. B **108**, 6571 (2004).
[31] N. Singhal, C. D. Snow, and V. S. Pande, J. Chem. Phys. **121**, 415 (2004).
[32] D. L. Mobley, J. D. Chodera, and K. A. Dill, J. Chem. Phys. **125**, 084902 (2006), following paper.
[33] W. C. Swope, J. W. Pitera, F. Suits *et al.*, J. Phys. Chem. B **108**, 6582 (2004).
[34] G. Jayachandran, V. Vishal, and V. S. Pande, J. Chem. Phys. **124**, 164902 (2006).
[35] H. Fujitani (private communication).