

HHS Public Access

Author manuscript

J Comput Chem. Author manuscript; available in PMC 2016 September 05.

Published in final edited form as:

J Comput Chem. 2015 September 5; 36(23): 1772-1785. doi:10.1002/jcc.23996.

Large Scale Asynchronous and Distributed Multi-Dimensional Replica Exchange Molecular Simulations and Efficiency Analysis

Junchao Xia^{1,*}, William F. Flynn^{1,2,*}, Emilio Gallicchio^{3,*}, Bin W. Zhang¹, Peng He¹, Zhiqiang Tan⁴, and Ronald M. Levy^{1,\$}

¹Center for Biophysics and Computational Biology, Department of Chemistry and Institute for Computational Molecular Science, Temple University, Philadelphia, PA 19122

²Department of Physics & Astronomy, Rutgers University, Piscataway, NJ 08854

³Department of Chemistry, CUNY Brooklyn College, Brooklyn, NY 11210

⁴Department of Statistics, Rutgers University, Piscataway, NJ 08854

Abstract

We describe methods to perform replica exchange molecular dynamics (REMD) simulations asynchronously (ASyncRE). The methods are designed to facilitate large scale REMD simulations on grid computing networks consisting of heterogeneous and distributed computing environments as well as on homogeneous high performance clusters. We have implemented these methods on NSF XSEDE clusters and BOINC distributed computing networks at Temple University, and Brooklyn College at CUNY. They are also being implemented on the IBM World Community Grid. To illustrate the methods we have performed extensive (more than 60 microseconds in aggregate) simulations for the beta-cyclodextrin-heptanoate host-guest system in the context of one and two dimensional ASyncRE and we used the results to estimate absolute binding free energies using the Binding Energy Distribution Analysis Method (BEDAM). We propose ways to improve the efficiency of REMD simulations: these include increasing the number of exchanges attempted after a specified MD period up to the fast exchange limit, and/or adjusting the MD period to allow sufficient internal relaxation within each thermodynamic state. Although ASyncRE simulations generally require long MD periods (> picoseconds) per replica exchange cycle to minimize the overhead imposed by heterogeneous computing networks, we found that it is possible to reach an efficiency similar to conventional synchronous REMD, by optimizing the combination of the MD period and the number of exchanges attempted per cycle.

Keywords

asynchronous replica exchange; molecular dynamics; distributed computing network; efficiency analysis; binding energy distribution analysis method; host-guest system

Correspondence to: ronlevy@temple.edu.

^{*}Equally contributed

Introduction

Molecular Dynamics (MD) simulations are widely employed to study the behavior of chemical and biological systems at the molecular level.¹⁻³ However, currently MD simulations are limited to time scales much shorter (< milliseconds) than those of many biochemical processes,⁴⁻⁶ even using the latest high performance computing resources or specialized computing chips.⁷⁻⁹ On the other hand, conformational equilibria of proteins and nucleic acids and the catalytic functions of enzymes and ribozymes often occur on time scales from milliseconds to seconds or longer.¹⁰ Besides utilizing more powerful computing hardware,^{5,7-9,11} developing more advanced conformational sampling techniques^{1-3,12-14} is an important alternative to address the timescale challenge of MD simulations. These enhanced sampling techniques are generally based on the imposition of thermodynamic or alchemical biasing forces on the relevant chemical reaction space and are able to speed up, often by many orders of magnitude, conformational interconversions otherwise too rare to be observed in traditional simulations.¹⁵⁻²⁷ Typically true (unbiased) thermodynamic observables have to be extracted from biased simulation results via postprocessing using reweighting techniques.²⁸⁻³⁶

Generalized ensemble methods³⁷⁻⁴³ are popular among the many enhanced conformational sampling methods and have been shown to provide better conformational mixing and faster convergence in many situations. These algorithms produce a random walk, not only in conformational space, but also in thermodynamic or Hamiltonian parameter spaces, such as the temperature of the system treated as a stochastic variable in the simulated tempering method. In a typical implementation, the information about the state of the system alternates between updates of particle positions and velocities from independent molecular dynamics (MD) or Monte Carlo (MC) simulations, and stochastic updates of thermodynamic conditions and/or Hamiltonian parameters (defined as the thermodynamic "state"), with the microscopic reversibility criteria applied to satisfy a valid canonical ensemble at each state. Generalized ensemble enhanced sampling implementations can be classified as either serial or parallel. In serial implementations (such as serial tempering and Hamiltonian hopping) only one MC/MD simulation thread is carried out in position space, and updates of the state of the system are performed periodically which requires iteratively adjusted free energy weights to equalize state populations visited.⁴⁴⁻⁴⁷ Since the determination of optimal free energy weights is equivalent to the computation of a free energy profile, it can be time consuming, especially when they are slowly convergent due to rare conformational transitions. In contrast, parallel replica exchange (RE) algorithms^{37-43,48,49} overcome the need of serial algorithms for the prior determination of free energy weights by launching many replicas (multiple independent MC/MD threads) at the same time. Those replicas are executed in parallel in such a way that there are as many replicas as thermodynamic states of the system included in the generalized ensemble and only one state is assigned to each replica. Periodically, replicas exchange their current state assignment with that of another replica, according to the probability of exchanges controlled by microscopic reversibility requirements for sampling the generalized ensemble spanning both the configurational space of each replica and the combinatorial set of assignments of states to replicas. The thermodynamic equivalency of replicas and the fact that there is always one replica at each

state, guarantee that at steady state each replica will visit each state with equal probability, a clear advantage over serial state-hopping algorithms dependent on prior knowledge of free energy weights.

This advantage of REMD is however counterbalanced by the need of RE for a parallel computational environment sufficiently large to host each replica of the system; a requirement that has historically discouraged the deployment of RE on a large scale. In our view, this is not necessarily due to the lack of availability of parallel computer hardware technologies-in recent years multi-core high performance computing clusters and computational grids have exponentially increased in both numbers and power-but rather to the lack of suitable software technologies capable of efficiently harnessing this latent computer power. Current implementations of the replica exchange method by the computational chemistry community are in fact severely limited in terms of its scalability and control when many replicas are involved. In conventional implementations of RE,³⁷⁻⁴³ simulations progress in unison and exchanges occur in a synchronous manner right after all replicas reach a pre-determined state (typically the completion of a certain number of MD steps, the MD period). This synchronous approach has several severe limitations. Firstly, sufficient dedicated computational resources must be secured for all of the replicas before the simulation can begin execution. Secondly, the computational resources must be statically maintained until the simulation is completed. Thirdly, a failure of any replica simulation typically causes the whole calculation to abort. Fourthly, the centralized synchronization prefers homogeneous computing environments otherwise the efficiency of Synchronous REMD (Sync REMD) will deteriorate due to the lag from the slowest computing unit. The reliance on a static pool of homogeneous computational resources and zero fault tolerance prevents the synchronous RE approach from being a feasible solution for new applications that demand multi-dimensional RE algorithms employing hundreds to thousands of replicas.⁵⁰⁻⁵² Besides the simulated tempering⁴⁷ and similar methods^{53,54} which require nontrivial pre-determined weighting factors, a multiplexed replica exchange method (MREM)⁵⁵ has also been proposed to perform RE simulations on the folding@home distributed computing environments⁵⁶ although exchanges between multiplexed replicas still require synchronization.47,55

In this work we introduce a replica exchange methodology named ASyncRE which removes the synchronizing concept. The basic idea of asynchronous RE is to assign all replicas to either the running or the waiting lists, and allowing a subset of replicas in the waiting list to perform exchanges independently from the other replicas on the running list. Because the exchanges do not rely on centralized synchronization steps, the ASyncRE algorithm is scalable to an arbitrary number of processors and avoids the requirement of maintaining a static pool of processors. Thus the method is suitable for deployment in both logically and physically distributed environments, in which the number of concurrently running replicas changes dynamically depending on the available resources. Prototypical implementations⁵⁷ have shown the potential range of benefits that can be achieved with dynamic execution and asynchronous RE. In addition to resiliency with respect to dynamically changing resources, we have shown that asynchronous RE also provides a number of important additional benefits, such as higher performance on clusters of machines with heterogeneous CPU speeds and the ability to employ complex exchange schemes that improve mixing by going

beyond conventional nearest neighbor communication. The challenge is to provide these capabilities at extreme scales, and with the flexibility and efficiency required to enable science applications currently out of reach. In this report we demonstrate how this challenge can be overcome by our recent implementations of the asynchronous RE algorithm which is capable of scaling to very large numbers of replicas and taking advantage of dynamically distributed and heterogeneous computational resources, including XSEDE high performance clusters, university grid networks consisting of spare computers on campus (Temple University and Brooklyn College at CUNY), and world-wide networks contributed by volunteer computing units (World Community Grid at IBM).

To improve the efficiency and convergence of conventional synchronous (Sync) REMD, many developments have been attempted, including modifying nonbonded potentials,⁵⁸⁻⁶⁰ simplifying the solvent contribution by solute tempering,⁶¹ graining solute structure,^{62,63} applying bias potentials, 27, 52, 64-66 performing exchanges with structure reservoirs, 67-69 and many others^{70,71} with Hamiltonian features. On the other hand, many investigations have focused on expanding exchange dimensions,^{39,42,52,72} building Markov state models,⁷³⁻⁷⁶ and optimizing the setting of simulation parameters such as the temperature distribution of the replicas, 37,77-83, the number of λ values, ⁷⁴ the exchange frequency, ⁸⁴⁻⁸⁷ and the number of exchanges attempted.^{73,88} The efficiency and convergence analysis of synchronous REMD in comparison with conventional MD has also been carried out in many previous studies.^{75,79,89-95} There is still debate concerning how to select simulation parameters such as the length of individual MD simulations (MD or exchange period) within a single cycle of MD + exchange, the number of exchanges attempted after an MD period within a single cycle, and the MD period when the total number of exchanges attempted and the total length of simulations are fixed. Some early results showed that the efficiency of REMD could be significantly reduced when the MD period is smaller than a certain number (1 ps).^{89,90,92} Recent results,^{84,85} however, found that the efficiency increases monotonically as the MD period becomes smaller and led to so called "infinite swap" methods.^{86,87,96} Previous results^{73,88,92,97} also illustrated that at regular MD periods, the number of attempted exchanges should be ideally chosen as large as feasible, namely increasing the number of exchanges within a single replica exchange cycle can improve the efficiency of REMD. No study can be found for "packing" the MD periods together with exchange attempts per cycle when the total number of exchanges attempted and the total length of simulation are fixed. Since ASyncRE simulations generally require large MD periods (> picoseconds) per RE cycle to minimize overhead from heterogeneous computing networks, and our file-based implementations of asynchronous (Async) REMD framework allows us more freedom to choose exchange settings, all of these become critical for achieving an efficient ASyncRE simulation protocol. Another difference between our implementations of Async and Sync REMD is that a replica in Async REMD simulations performs exchanges with all other replicas not limited to its nearest neighbors in thermodynamic space as is the case for traditional Sync REMD simulations. One of the goals of this report is to present the relevant efficiency analysis in the context of the ASyncRE methodology.

Methods

Replica Exchange Sampling

The conformational sampling problem can be formalized as the problem of efficiently drawing samples of molecular conformations x from the canonical distribution of the chemical system:

$$p(x|\beta,\theta) = \frac{exp\left[-\beta U(x|\theta)\right]}{Z(\beta,\theta)}, \quad (1)$$

where $U(x|\theta)$ is the potential energy function of the system at molecular configuration *x*, parametrically dependent on environmental conditions (volume, etc.), chemical composition (molecular topologies, concentrations, etc.) and modeling parameters (partial charges, QM basis functions, biasing potential settings, etc.) collectively denoted as θ . For the following we will define the dimensionless potential energy function $u(x|\beta, \theta) = \beta U(x|\theta)$, which depends parametrically on both the inverse temperature β as well as the system parameters θ . For ease of notation we will denote the *state* of the system as $s = (\beta, \theta)$, fully specified by the joint set of inverse temperature and system parameters. $Z(\beta, \theta) = Z(s)$ is the canonical configurational partition function at state $s = (\beta, \theta)$ defined such that Eq. (1) is normalized with respect to *x*. Metropolis Monte Carlo (MC) and Molecular Dynamics (MD) are two standard molecular modeling methods to sample Eq. (1). These however are limited by slow equilibration rates due to rarely crossed energy barriers and entropic bottlenecks connecting stable conformational domains.

The Replica Exchange (RE) method^{37,38} attempts to enhance sampling by considering the extended ensemble described by the distribution

$$p_{RE}(x_1, x_2, \dots, x_M | \{s\}) = exp \left[-\sum_{i=1}^{M} u(x_i; s[i]) \right] / Z_{RE} \quad (2)$$

where the index *i* denotes one of *M* realizations of the system, called *replicas*, with molecular configuration x_i at the assigned state s[i] taken from a discrete set $(s_1, s_2, ..., s_M)$ of *M* possible states without repetition, such that no state is assigned to more than one replica (although equivalent states in the state set are allowed). The symbol $\{s\}$ denotes one of the *M*! permutations of the assignment of states to replicas and s[i] is the state assigned to replica *i* according to the given permutation. For example with three replicas and three states, $(s[1] = s_2, s[2] = s_1, s[3] = s_3)$ is one such permutation, in which state s_2 is assigned to replica 1, state s_1 is assigned to replica 2, and state s_3 is assigned to replica 3.

Because there are no cross terms, the partition function, Z_{RE} , corresponding to the normalization factor in Eq. (2), is given by the product of the partition functions at each state

$$Z_{RE} = Z(s_1) Z(s_2) \dots Z(s_M) \quad (3)$$

and is, consequently, independent of the state permutation $\{s\}$. It follows that any thermodynamic quantity of the replica exchange extended ensemble can be computed with an arbitrary state permutation, or that, equivalently, any two permutations will result in the

same value of thermodynamic quantities. This property is exploited in the *replica exchange conformational sampling* method in which replicas are allowed to explore both conformational space, x, and chemical/parameter space s by sampling the discrete space of M! state permutations {s}. Formally the method samples the joint distribution

$$p_{RE}(x_1, x_2, \dots, x_M, \{s\}) = p(\{s\}) = p_{RE}(x_1, x_2, \dots, x_M | \{s\}) \propto p_{RE}(x_1, x_2, \dots, x_M | \{s\}), \quad (4)$$

where $p_{\text{RE}}(x_1, x_2, ..., x_M | \{s\})$ is defined above and we have chosen uniform prior probabilities, $p(\{s\}) = 1/M!$, of state permutations. Similarly, Eq. (4) can be equivalently written as

$$p_{RE}(x_1, x_2, \dots, x_M, \{s\}) = p(x_1, x_2, \dots, x_M) p_{RE}(\{s\} | x_1, x_2, \dots, x_M) \propto p_{RE}(\{s\} | x_1, x_2, \dots, x_M\}), \quad (5)$$

where we have assumed uniform probability, $p(x_1, x_2, ..., x_M)$, in configurational space in absence of potential energy. Comparing Eqs. (4) and (5) we see that the conditional probability $p_{\text{RE}}(x_1, x_2, ..., x_M | \{s\})$ of molecular configurations given the permutation of states and the conditional probability $p_{\text{RE}}(\{s\}|x_1, x_2, ..., x_M\})$ of state permutations are given by the same expression [the numerator of Eq. (2)] by interpreting it alternatively in terms of $(x_1, x_2, ..., x_M)$ or $\{s\}$ as the independent variables.

Asynchronous Replica Exchange

In conventional *synchronous* implementations of RE^{37,38} the reassignment of states to replicas is coordinated by a master process (typically implemented using MPI) and occurs simultaneously for all replicas after these have reached a suitable synchronization point, such as the completion of a given number of MD steps (MD period). Synchronous RE (Sync RE) is a suitable algorithm for stable, tightly coupled, and uniform computing architectures, such as a large High Performance Computing (HPC) cluster, where many MD threads can efficiently execute in parallel at equal speeds for extended periods of time without failures. When these conditions are met, it is straightforward to implement synchronous RE algorithms capable of achieving a high rate of exchanges with minimal impact of the MD stoppage time on the overall throughput.

Synchronous implementations of RE, however, are either not feasible or extremely inefficient in heterogeneous environments, as in the extreme case of volunteered computational grids such as IBM's World Community Grid (WCG). In these environments interprocess communication across compute nodes is typically not available, and the pool of compute nodes changes dynamically without guarantee of stability or homogeneity. Similar concerns exist for larger installations as, for example, when attempting to straddle one large coupled parallel simulation across two or more HPC clusters connected by a thin network link. As illustrated in this work, for multi-dimensional RE simulations involving a large number of replicas (hundreds to thousands), there are clear benefits of alternatives to synchronous RE in terms of resource allocation, resiliency to failure, and ease of implementation even on tightly coupled HPC clusters.

Unlike parallel numerical algorithms, such as molecular dynamics, requiring synchronization between parallel threads, the replica exchange method itself does not

impose the restriction that exchanges should necessarily occur synchronously across all processors. In particular the RE method itself does not require that all of the replicas be running at the same time. There are therefore no obstacles in principle preventing the deployment of RE over distributed and heterogeneous computing infrastructures. An asynchronous RE algorithm based on a decentralized over-the-network mechanism has been developed by some of us some years ago.^{57,98} That work showed that the asynchronous prescription can provide significant advantages over the conventional synchronous implementation in terms of scalability with increasing number of replicas, both with respect to CPU utilization and the ability to employ a non-nearest neighbor exchange scheme leading to improved mixing in configurational and state spaces.

In this work we propose a similar algorithm in spirit but based on a coordination server that conducts exchanges on the file system where replicas not currently running are checkpointed. The algorithm can be described schematically as follows:

- **1.** Job files and executables for each replica are set up locally as appropriate depending on the application.
- 2. Periodically, a subset of the replicas are submitted for execution of MD simulation remotely. At the same time the output of remote replicas that have completed an MD execution cycle are collected.
- **3.** Periodically, exchanges of thermodynamic parameters are performed between the local replicas not currently executing. The energetic and structural information required for the exchange steps are collected from the output files of the replicas. Swaps are implemented by replacing values of parameters, in the MD engine input files as appropriate. New cycles are then initiated by re-submitting replicas for execution (point 2).

It is evident that in this algorithm exchanges occur asynchronously, that is for example they occur for some replicas while other replicas are undergoing MD. The algorithm does not require a direct network link between the compute nodes as all exchanges occur on the file-system of the coordination server. Furthermore the algorithm does not rely on a static pool of compute nodes as each run cycle of a replica can occur on a different compute node that does not need to be secured in advance. We have implemented the ASyncRE methodology for XSEDE high performance cluster resources, the BOINC distributed computing for campus grid networks like the ones at Temple University and Brooklyn College at the City University of New York, and we are working on an implementation for the world-wide distributed BOINC networks like World Community Grid (WCG) at IBM consisting of 650,000 volunteers and 2,700,000 computing units. A brief introduction to the specific implementations is included as an Appendix and a more complete description of the software will be published soon.⁹⁹ The software is free to download at (https://github.com/ComputationalBiophysicsCollaborative/AsyncRE).

Replica Exchange Scheme

In RE, sampling is performed through a Markov chain alternating between updates of molecular configurations using MC or MD independently for each replica at a fixed state,

and updates of state assignments to replicas (permutations) via a series of coordinated attempted swaps of states among pairs of replicas according to MC algorithms. In the simplest variation, a new permutation of states $\{s'\}$ is proposed at random from the current permutation $\{s\}$ typically by swapping two randomly picked indexes, corresponding to exchanging states among a pair of replicas. The proposal is then accepted with probability

$$p_{ss'} = min\left\{1, \frac{p_{RE}\left(\left\{s'\right\}|x_{1}, x_{2}, \dots, x_{M}\right)}{p_{RE}\left(\left\{s\right\}|x_{1}, x_{2}, \dots, x_{M}\right)}\right\} = min\left\{1, exp\left[\sum_{i=1}^{M} u\left(x_{i}; s\left[i\right]\right) - \sum_{i=1}^{M} u\left(x_{i}; s'\left[i\right]\right)\right]\right\}$$
(6)

based on the well-known Metropolis scheme. The acceptance probability of this randomly picked update is typically very low unless it is downhill (such as, for example, when giving the lower temperature to the replica with the lower potential energy) or when the permutation involves neighboring states. For large multi-dimensional simulations covering a large portion of state space the probability of picking a pair of replicas with neighboring states can be very small. Because of the high rate of rejections this algorithm (referred to as "Metropolis all-to-all") results in slow diffusion in parameter space.

One alternative (referred to as "Metropolis nearest-neighbor") is to select pairs of replicas for exchange so as to minimize the rejection probability of the exchange. In conventional one-dimensional implementations of RE (those in which only the temperature or a single thermodynamic parameter is varied at constant temperature) it is common to limit exchanges between neighboring replicas, that is those that hold immediately adjacent states, (e.g. in the nearest neighbor exchange scheme a set of attempted exchanges is performed between paired nearest neighbor thermodynamic states). On the other hand, for more complex multidimensional RE implementations, limiting attempted exchanges to neighboring states can be problematic. For the ASyncRE methodology we designed, only the replicas in the waiting list can participate in the exchange process, therefore we implemented an algorithm for the sampling of the state permutation space which does not require the prior identification of neighboring states, similar to the Metropolis-based independence sampling (MIS) algorithm by Chodera & Shirts,⁸⁸ which attempts to exchange two replicas randomly picked but follows the same Metropolis criterion, and it was postulated that this algorithm approaches the Gibbs sampling limit when the number of swaps is of the order of M^3 to M^5 .

BEDAM Method and UWHAM Reweighting for Estimating Absolute Binding Free Energy

The Binding Energy Distribution Analysis Method (BEDAM),¹⁰⁰⁻¹⁰⁴ a novel approach for absolute binding free energy estimation and analysis developed in our group, is based on a sound statistical mechanics theory¹⁰⁵ of molecular association and efficient computational strategies built upon parallel Hamiltonian replica exchange sampling (λ hopping)^{61,106} and thermodynamic reweighting.^{28,31,33,35} The total potential energy of the receptor-ligand complex can be reduced to the dimensionless form as

$$u(x;\beta,\lambda) = \beta \left[U_0(x) + \lambda b(x) \right] \quad (7)$$

where λ is an alchemical progress parameter ranging from 0, corresponding to the uncoupled state of the complex, to 1, corresponding to the fully coupled state of the complex. $U_0(x)$ is

the potential energy of the complex when receptor and ligand are uncoupled, that is as if they were separated at infinite distance from each other. The quantity b(x), called the binding energy, is defined as the change in effective potential energy of the complex for bringing the receptor and ligand from infinite separation to the given conformation x of the complex.

The BEDAM method calculates the binding free energy ΔG_b° between a receptor A and a ligand B using the AGBNP implicit solvation model^{107,108} as :

$$\Delta G_b^{\circ} = -k_B T \ln \left[C^{\circ} V_{site} \int db p_0 (b) e^{-\beta b} \right] = -k_B T \ln C^{\circ} V_{site} + \Delta G_b, \quad (8)$$

where $\beta = 1/k_BT$, C° (=1M) is the the standard concentration of ligand molecules, V_{site} is the volume of the binding site, and $p_0(b)$ is the probability distribution of binding energy (b(x) in Eqs. 7 and 8) collected in an appropriate decoupled ensemble of conformations in which the ligand is confined in the binding site while the receptor and the ligand are not interacting with each other but both only with the solvent continuum.

Earlier versions of BEDAM¹⁰⁰ were implemented in our IMPACT¹⁰⁹ molecular simulation package using the synchronous nearest-neighbor exchange scheme and performing the Hamiltonian replica exchange only in the λ space. Diffusion along the λ variable connects the bound and unbound conformational states and accelerates the exploration of intermolecular degrees of freedom. The ability to carry out extensive conformational sampling of relative position and orientation of the ligand with respect to the receptor is an advantage of BEDAM¹⁰⁰⁻¹⁰⁴ over existing free energy perturbation (FEP) and absolute binding free energies protocols in explicit solvent.

In the multi-dimensional BEDAM method as illustrated in Fig. 1 and implemented as the ASyncRE framework, the system is modeled with λ and β as exchange parameters. The purpose of sampling along λ is to enhance mixing of conformations along the alchemical pathway while high temperatures enhance sampling of internal molecular degrees of freedom at each alchemical state. We mention that the temperature is only one of the additional thermodynamic coordinates designed to activate more thoroughly intramolecular degrees of freedom that can be implemented in our file-based ASyncRE implementations using the IMPACT MD engine. In Figs. S1 and S2 of Supporting Information, we show several time series of λ and temperature for replicas from the 1D and 2D Async RE simulations. It is clear that in both 1D and 2D Async RE implementations the diffusion speed in thermodynamic state space increases as the number of exchanges is increased after a MD period.

In this work, we employed the unbinned weighting histogram analysis method $(\text{UWHAM})^{34,35}$ to estimate binding energy distributions $p_0(b)$ and binding free energies

 $\Delta G_b = -\frac{1}{\beta} \int p_0 (b) e^{-\beta b} db$ from binding energy samples obtained from the HREM simulations. The unbinned WHAM is a recent development and can treat, in a very efficient way, sparsely distributed interaction energy samples as obtained from unmodified interaction potentials that are difficult to analyze using standard binning methods.^{28,30} The

theoretical derivation and numerical validation of UWHAM have been reported recently,^{34,35} along with the comparison with other methods such as WHAM^{28,30} and the multi-state Bennet acceptance ratio method (MBAR).³³ The code implementation also has been incorporated into the R package available at (http://cran.r-project.org/web/packages/UWHAM/index.html).

Metrics of Efficiency and Convergence

Currently there is no universally accepted metric for evaluating the sampling efficiency in REMD simulations although several methods have been utilized in previous work such as the slowest eigenvalue of the Markov chain from an analysis of the state transition matrix,^{79,88} correlation time and end-to end transit of the replica state index,⁸⁸ variances of the estimated means of relevant observables,^{75,91,93} and root-mean-square deviation (RMSD) of related observables of test simulations from the corresponding reference simulation.^{84,85} In this work, we performed statistical inefficiency analysis¹¹⁰ and extracted the total effective relaxation time of the binding energy at $\lambda = 1.0$, which includes all relaxation effects both from MD simulations and replica exchange mixing. The statistical inefficiency (*s*) can be extracted from the block averaging of the binding energy series and is related to total effective relaxation time of binding energy (τ_u) as below, ¹¹⁰

$$s = \lim_{T_b \to \infty} \frac{T_b \, \sigma^2 \, (\langle u \rangle_b)}{\sigma^2 \, (u)} \approx 2\tau_u, \quad (9)$$

where the whole series of length T is divided into n_b blocks of length T_b and $\langle u \rangle_b$ is the

 $\sigma^{2}(\langle u \rangle_{b}) = \frac{1}{n_{b}} \sum_{b=1}^{n_{b}} (\langle u \rangle_{b} - \langle u \rangle)^{2}$ is the corresponding variance from n_{b} blocks. In contrast, $\langle u \rangle$ and $\sigma^{2}(u)$ are the average and variance of the binding energy calculated from the whole series $T = T_{b}n_{b}$.

To evaluate the divergence of the binding energy distribution from a target distribution, we also calculated the Kullback-Leibler (KL) divergence¹¹¹ defined as

$$D_{KL}(P||Q) = \sum_{i} P(i) \ln \frac{P(i)}{Q(i)}, \quad (10)$$

which represents an approximated distance between the calculated distribution Q(i) and the target one P(i). Namely $D_{KL} \rightarrow 0$ as $Q(i) \rightarrow P(i)$ for all bins from the distributions.

Results

1D Sync REMD Simulations

The model complex we studied is b-cyclodextrin-heptanoate as depicted in Fig. 2, one of the host-guest systems investigated in our previous work.^{101,103} As a benchmark we performed standard 1D Synchronous BEDAM simulations at 16 λ values (0.0, 0.001, 0.002, 0.004,

0.01, 0.04, 0.07, 0.1, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 0.95, and 1.0) and two different temperatures (200 and 300K respectively) with a MD period of 0.5 ps per RE cycle, using the OPSL-AA force field^{112,113} and the AGBNP2 implicit solvent model.¹⁰⁸ The binding energy distributions at $\lambda = 1.0$ are displayed in Fig. 2 (the results shown correspond to the distributions obtained after UWHAMing) for both 200K and 300K, and were calculated from 1.152 μ s aggregated simulations with 16 replicas (72 ns for each replica). The multiple peaks in the binding energy distributions correspond to different orientations of heptanoate in the host cavity, each characterized by different hydrogen bonding patterns with bcyclodextrin.^{101,103} The binding free energies estimated from these two simulations are -0.60 and -6.58 kcal/mol for 300 and 200K respectively. These two long 1D Sync REMD simulations will serve as the golden standard for evaluating our 1D and 2D Async REMD simulations with different combinations of MD simulation period per RE cycle and the number of exchanges attempted after the MD period per RE cycle. All 2D Async REMD simulations have the same 16 λ values and the temperature dimension is extended to the following fifteen temperatures (200, 206, 212, 218, 225, 231, 238, 245, 252, 260, 267, 275, 283, 291, 300K), resulting in a total of $16 \times 15 = 240$ replicas with 240 different pairs of (λ , β) values.

Convergence and Efficiency Analysis for 1D Async REMD

Figure 3a shows the binding free energies as a function of simulation time from 1D Async REMD using different combinations of the MD period per RE cycle and the number of exchanges attempted per RE cycle, along with the golden standard from 1D Sync REMD. The binding free energies from all simulations converge to a value of around -0.60 kcal/mol at 300K before 30ns. The final distributions of binding energies from 1D Asnyc REMD are consistent with that of Sync REMD as displayed in Fig. 3b and also in Fig. 3c for KL divergence curves. The slight differences around the three peaks originates from the accumulated effects of the different initial conditions since all data points in the time series including the initial equilibration have been included in the distributions.

Figure 3d shows the results of the block averaging of binding energies from 1D Async REMD simulations when only one exchange was attempted per RE cycle after an MD period (1ps or 10ps). From the values of inefficiency *s* (the plateau value when the block size increases), it is clear that for a fixed number of exchange attempts per RE cycle a shorter MD period is more efficient, which is consistent with previous work from other groups.^{84,85,89,90,92} The underlying rationale is that the REMD simulation with a longer MD period involves a smaller number of total exchanges attempted per unit time, and therefore results in a slower equilibration of replicas. This observation suggests an apparent drawback of asynchronous implementations of replica exchange since the MD period generally can not be too small (> 1ps) due to the overhead from the latency response of the local filesystem and/or the allocating and transferring of MD jobs by the job manager. But this is not really a problem, see the following results and discussion below.

When the MD period per cycle is fixed to a large value (10 ps) as in Fig. 3e, the REMD simulation becomes more efficient as the number of attempted exchanges is increased (from 10 to 100). This observation is in agreement with our previous studies of Markov

models^{73,97} of replica exchange and also REMD simulations by other groups.^{88,92} The fact that simulations with 50 and 100 exchange attempts have almost the same efficiency suggests that a fast exchange limit exists when the MD period per cycle is held constant. We explore the issue of the fast exchange limit further in related work.¹¹⁴ From Fig. 3e, it can also be seen that the inefficiency *s* values of Async REMD simulations (with 50 or 100 exchanges) can be smaller than that of Sync REMD simulation with a typical MD period of 0.5ps. This information is critical for improving the efficiency of Async REMD simulations where the overhead from heterogeneous and distributed computing environments necessitates longer MD periods per RE cycle (> 1ps). Namely, we can increase the number of exchanges per RE cycle, until the fast exchange limit is reached, and reduce the replica mixing time (improve the efficiency) to a value comparable to the Sync REMD benchmark (which employed one set of nearest-neighbor exchange per cycle) despite the longer MD period required for Async REMD.

Moreover, from Fig. 3f, we also found that the efficiency increases as the MD period is *increased* to 10ps from 1ps when the number of exchanges attempted per RE cycle is also increased, keeping the ratio of exchange attempts per cycle to the MD period per cycle constant (fixed at 1 or 5 exchanges per ps in Fig. 3f). This might be considered a surprising result since a fixed ratio implies that the total number of exchanges attempted is the same for the same total simulation length. However these can be "packed" using different MD periods per cycle, such as 1ps MD and 1 exchange, and 10ps MD and 10 exchanges. Figure 3f illustrates that the longer MD period can lead to higher efficiency when the total number of attempted exchanges per unit time is fixed, suggesting that a minimum value of the MD period may be required to allow sufficient internal relaxation at a thermodynamic state and improve efficiency when the simulations have the same replica mixing effects (the same total number of attempted exchanges per unit time). We return to this subject in the discussion.

The efficiency analysis above is based on binding energies which is the most direct and relevant quantity for binding free energy calculations. We note that similar conclusions are also valid for other quantities related to conformational changes. In Fig. S3 of Supporting Information, we show the results using the orientational angles of the heptanoate relative to the ring plane of cyclodexin, which mainly determines the binding poses of the ligand. The results are consistent with Figs. 3e and 3f, although in general the derived total effective relaxation times of orientational angles are larger than that from binding energies.

Convergence and Efficiency Analysis for 2D Async REMD

Multi-dimensional REMD is another way to improve the efficiency of REMD simulations.^{39,42,52,115} We have extended the 1D Async REMD simulations (16 λ values) to 2D Aysnc REMD simulations using the additional temperature space (15 values), resulting in 240 replicas and 7.2 μ s aggregated simulation time (30 ns for each replica). Figures 4a and 4b show the final binding energy distributions and corresponding KL divergences at λ = 1.0 and T=300K calculated from several 2D Async REMD simulations with different combinations of the MD period and the number of exchanges attempted per RE cycle. All

binding energy distributions converge to the result of the standard 1D Sync REMD as in the case of 1D Async REMD but in a shorter simulation time (about 20ns versus 30ns).

The efficiency results in Fig. 4 suggest similar trends as that of 1D Async REMD. The 2D Async REMD simulation becomes more efficient as the number of exchanges attempted per RE cycle is increased when the MD period is fixed (10ps Fig. 4c); efficiency is also improved as the MD period is increased to 10ps from 1ps when the ratio of the number of exchanges attempted to the MD period is fixed at 1 per ps (Fig. 4d). We also note that the s value of the most efficient parameter choice (10ps for the MD period and 8000 for the number of exchanges attempted per RE cycle) for 2D Async REMD is almost half of that of 1D Async REMD (Fig. 3e), indicating that the efficiency of 2D REMD in the fast exchange limit can be at least twice as good as running 1D REMD with a standard choice of parameters. It is obvious that 2D REMD simulations require much more computing resources due to the extension of the temperature dimension, but the simulations also provide binding free energies at different temperatures (see Figs. 4e and 4f for results at T=200K). More importantly, 2D REMD simulations require much less computing time per thermodynamic state due to the faster convergence than that of 1D REMD, since hightemperature replicas can accelerate the sampling of low-temperature replicas. We note that in this work, we performed the 2D REMD in a temperature range below 300K (200, 206, 212, 218, 225, 231, 238, 245, 252, 260, 267, 275, 283, 291, and 300K), in order to mimic stronger binding free energies which are comparable in strength to values typically observed in protein-ligand systems.

Efficiency of 2D Async REMD on a BOINC Distributed Network

The efficiency results above were obtained from our Async REMD implementation for XSEDE high performance resources and only one homogeneous cluster (Gordon, Trestles, or Stampede) was involved for each set of Async REMD simulations. In those simulations all replicas have the same CPU wall time for all individual MD simulations. In contrast, for the BOINC distributed network at Temple University consisting of 450 CPUs in teaching laboratories, the CPU wall times for individual MD simulations (100 ps for the b-cyclodextrin-heptanoate complex) have wide distributions as shown in Fig. 5a. The multimode distribution is not only due to the heterogeneous hardware resources, but also to the different interrupt patterns of usage, since the BOINC clients are set so that a running MD simulation will be halted once a user login is detected and will be resumed after logout. This heterogeneity of the BOINC distributed network also resulted in a wide distribution of the number of exchanges attempted, displayed in Fig. 5b (with a mean value around 1200 exchanges in a 100ps MD period) across the 240 replicas. As such, for the BOINC benchmarks we can only specify an estimated wall clock time of exchange cycling for all replicas.

One limitation of Async REMD is that the MD period can not be too short as illustrated in Table 1 in order to reduce the overhead inherent in the BOINC management infrastructure including preparing input files in the local filesystem, delayed queuing inherent to the BOINC server, and submitting input files to and receiving output files from remote BOINC clients. We can see that, from Table 1, the fraction of overhead can be cut down greatly as

the MD period is increased from 10 ps to 100 ps for the b-cyclodextrin-heptanoate complex, or as the system size becomes large enough, comparable to that of protein-ligand systems. Shown as an example in Table 1 is the timing for simulations of the enzyme ABL kinase. We note that the fraction of overhead can be minimized by optimizing the MD period through a series of ASync REMD simulations at different MD periods. However, this trade-off value of MD period, balancing the fraction of time spent on overhead which goes up as the MD period decreases and the efficiency of Async RE algorithm which becomes more efficient as the MD period decreases, depends on the system size and other settings of distributed networks.

Due to the inefficiency of Async REMD if the MD period is too short, we are interested in analyzing how efficient the Async REMD can be when a large MD period (such as 100 ps) is selected. Figure 5c shows the binding energy distribution calculated from a 2D Async REMD simulation using the Temple BOINC grid network, which is in agreement with 1D and 2D Async REMD simulations using XSEDE resources. The statistical inefficiency value from this simulation is slightly larger than the standard 1D Sync REMD as shown in Fig. 5d.

We point out that the number of exchanges attempted in this case is only an average value (1200) from a wide distribution (see Fig. 5b) due to the heterogeneous nature of the BOINC distribution network. The value of inefficiency *s*, however, can be decreased greatly by setting a much smaller wall clock time for exchange cycling to increase the average number of exchanges attempted. Hence the efficiency of 2D REMD using BOINC shown in Fig. 5d is not yet saturated by the fast exchange limit and can be improved further.

Discussion

The ASyncRE methodology described in this work, as illustrated by the applications, is capable of supporting large-scale and flexible execution of replica exchange calculations with hundreds of replicas. The algorithm is flexible in the sense that it supports different coupling schemes between the replicas. The basic idea behind the design of ASyncRE is to allow pairs of replicas to perform exchanges independently from the other replicas. Because it does not rely on centralized synchronization steps, the algorithm is scalable to a very large number of processors and avoids the requirement of maintaining a static pool of processors. Thus the method is suitable for deployment in both logically and physically distributed heterogeneous environments, in which the number of concurrently running replicas changes dynamically depending on the available resources. In contrast, synchronous replica exchange is designed for use on stable, tightly coupled, and homogeneous computing architectures, such as a large High Performance Computing (HPC) clusters, where many MD threads can efficiently execute in parallel at equal speeds for extended periods of time without failures. For simplicity in performing our efficiency analysis, we set the number of replicas on the waiting list to be roughly equal to the number on the running list. It should be possible to optimize the ratio between the number on the waiting list and that on the running list since intuitively more replicas on the waiting list will improve the exchange efficiency of thermodynamic states but reduce the number of MD simulations completed in a fixed clock time and slow down the diffusion speed in conformational space during the MD steps. So there is a tradeoff here which we will investigate in a future communication.

The major drawback of the file-based AsyncRE methodology is that the MD period has to be large enough in order to reduce the fraction of overhead due to the preparation of local files required to launch the MD simulations every cycle, the job queuing of the BOINC server, and the file transferring to and from remote BOINC clients. However, our results also show that the loss of efficiency resulting from longer MD periods per cycle can be compensated by increasing the number of exchanges attempted per cycle up to the fast exchange limit. In addition, the selected MD period can also to be shorter as the system size is increased, as shown in the case of ABL kinase. Moreover, the ASyncRE framework is very robust to failures of individual MD processes since no synchronizing process is required and failed jobs can be resubmitted automatically by the job manager.

The replica exchange MC algorithm implemented in our ASyncRE is based on randomly picking exchange pairs, so called Metropolis-based independence sampling.⁸⁸ Our results for 1D and 2D Async REMD simulations show that increasing the number of exchanges attempted per replica exchange cycle can significantly reduce the total effective relaxation time of the binding energy and improve the efficiency of REMD simulations. However, for the traditional implementation of REMD, the exchange process is synchronized after all individual MD simulations have finished. The total wall clock time for exchanges will be increased significantly if the completion of exchanges requires recalculating the potential energies in the new states (in the dimensions of the atomic positions (x) and the thermodynamic parameters (β , λ)) through the energy functions in the MD code. In our case, the potential energy $u(x, \beta, \lambda)$ can be decomposed into a linear combination of $u_0(x)$, b(x), β and λ (see Eq. 7). The reevaluation of new energies can be processed locally in a very fast way because it only involves recombination of these four terms from the output of MD simulations and does not need to call MD energy functions remotely. However, for some systems using non-linear soft-core potentials for ligand binding, the potential energy cannot be decomposed linearly and MD energy functions have to be recalled to obtain the energies of new states combining the information of new thermodynamic parameters and atomic positions. For those cases, the wall clock time for reevaluating new energies for exchanges may become comparable to the time required for the MD simulations as the number of attempted exchanges becomes large.

Through our experiments with 1D and 2D Async REMD simulations, we found three possible ways to improve the efficiency: (a) reduce the MD period per cycle when the total number of exchanges attempted per cycle is fixed; (b) increase the number of exchanges per MD period to reach the fast exchange limit when the MD period is fixed; (c) adjust the MD period so that it is not smaller than a minimum value which allows for sufficient internal relaxation while adjusting the number of exchanges attempted per cycle so as to remain in the fast exchange limit. Determining the most optimal set of parameters can be challenging since the fast exchange limit and the minimum value of the MD period are not a priori known and can vary from system to system. For the b-cyclodextrin-heptanoate host-guest system, the minimum value of the MD period is between 1 and 10ps, and the fast exchange limit for the MD period of 10 ps is 10 to 100 exchange attempts per replica and per exchange cycle. There are some signs that these values can be related to the roughness of the

energy landscape, and we are working on a more theoretical analysis of this problem using Markov state models to simulate the replica exchange process.¹¹⁴

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work has been supported by the National Science Foundation (CDI type II 1125332) and the National Institutes of Health (GM30580 and P50 GM103368). E.G. acknowledges support from the National Science Foundation (SI2-SSE 1440665). REMD simulations were carried out on the Gordon, Trestles, and Stampede clusters of XSEDE resources (supported by TG-MCB100145 and TG-MCB140124), and BOINC distributed networks at Temple University and Brooklyn College of the City University of New York. The authors acknowledge the great support from Gene Mayro, Jaykeen Holt, Zachary Hanson-Hart from the IT department at Temple University, and Sade Samlalsingh, James Roman, and John Stephen at Brooklyn College.

Appendix

ASyncRE Software Framework, Modular Design, and Implementations on Different Computer Resources

As discussed above, the reliance on a static pool of computational resources and zero fault tolerance prevents the synchronous RE approach from being a feasible solution for new application areas that demand multi-dimensional RE algorithms employing hundreds to thousands of replicas. To address these challenges we have developed a novel Python package named ASyncRE for distributed replica-exchange applications (https://github.com/ ComputationalBiophysicsCollaborative/AsyncRE). An upcoming publication will focus on the design of the software. Briefly, as illustrated in Fig. 6, the idea behind ASyncRE is the implementation of replicas as independent executions of the MD engine for a predetermined amount of simulation time. Each replica lives in a separate sub-directory of a local coordination server where the ASyncRE application runs. MD engine input files are prepared for each replica according to the RE scheme under consideration. As resources become available, a randomly chosen subset of the replicas are submitted to a Job Manager, which launches them on remote resources using a direct ssh link or through a BOINC infrastructure, and enter a running state. When a replica completes a cycle remotely (for example on XSEDE compute nodes or a BOINC client), the output data is transferred back to the server and the replica enters a waiting state, making it eligible for exchange with other replicas as well as the initiation of a new cycle. Periodically, exchanges of thermodynamic parameters are attempted between replicas in a waiting state using the Metropolis Independent Sampling algorithm⁸⁸ described above restricted to only the pool of replicas in the waiting state. Exchanges are conducted based on the appropriate reduced energies as specified in user-defined modules. This usually entails, see below, extracting energetic and structural information from the MD engine output files. Exchanges result in a new set of MD engine input files ready to begin a new execution cycle.

The ASyncRE software is modularized by taking advantage of the object-oriented capabilities of the Python language (class inheritance and method overrides), including three major components as described below.

- 1. modules to interface MD engines. These modules facilitate the interaction with the specified MD engine (IMPACT in this case), such as providing routines to compose input control files, and to read output files for collecting MD simulation results.
- modules to perform common tasks such as job staging through job manager and 2. coordinating exchanges of parameters among replicas. Independence sampling algorithms are implemented in the core module, which often calls specialized routines defined in user-provided modules implementing specific RE schemes with a given MD engine (temperature, Hamiltonian, etc. including multidimensional combinations of these). Currently modules for multidimensional RE and BEDAM λ -RE alchemical binding free energy calculations with the IMPACT MD engine are provided. One key function of modules implementing RE schemes is the computation of the reduced potential energy matrix $u_{ii} = u(x_i; s_i)$ in Eq. (2), containing the reduced potential energy of each replica i at each of the M states s_i . This, and the list of waiting replicas, is the only input for the independence sampling exchange algorithm implemented in the core module. RE modules also often override generic input/output routines in the MD engine modules to, for example, extract specific energetic information from output files to compute the reduced potential energy matrix.
- 3. modules to utilize different job transport mechanisms. An early design and initial usage of the ASyncRE software (https://github.com/saga-project/asyncre-bigjob) is described in a recent report.⁹⁸ To hide most of the complexities of resource allocation and job scheduling on a variety of architectures from large national supercomputing clusters to local departmental resources, we have recently implemented two different job transport systems: SSH transport for high performance cluster resources (such as those of XSEDE), and the BOINC transport for distributed computing on campus grid networks like the ones at Temple University and Brooklyn College at the City University of New York. Our group also received an invitation to join the FightAIDS@home project (http://fightaidsathome.scripps.edu) and will have access to the IBM computing resources, World Community Grid (WCG), a distributed BOINC volunteer grid network in a much larger scale (650,000 volunteers, 2,700,000 computing units).

References

- 1. Gallicchio E, Levy RM. Curr Opin Struct Biol. 2011; 21:161–166. [PubMed: 21339062]
- 2. Gallicchio E, Levy RM. Adv Prot Chem Struct Biol. 2011; 85:27-80.
- Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Curr Opin Struct Biol. 2011; 21:150–160. [PubMed: 21349700]
- 4. Zwier MC, Chong LT. Curr Opin Pharmacol. 2010; 10:745-752. [PubMed: 20934381]
- Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. Science. 2010; 330:341–346. [PubMed: 20947758]

- Lane TJ, Shukla D, Beauchamp KA, Pande VS. Curr Opin Struct Biol. 2013; 23:58–65. [PubMed: 23237705]
- 7. Bowers KJ, Dror RO, Shaw DE. J Comput Phys. 2007; 221:303–329.
- Bowers, K.; Chow, E.; Xu, H.; Dror, R.; Eastwood, M.; Gregersen, B.; Klepeis, J.; Kolossváry, I.; Moraes, M.; Sacerdoti, F.; Salmon, J.; Shan, Y.; Shaw, D. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. Proceedings of the ACM/IEEE Conference on Supercomputing (SC06); Tampa, Florida. 2006.
- 9. Güetz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. J Chem Theory Comput. 2012; 8:1542–1555. [PubMed: 22582031]
- 10. Boehr DD, Nussinov R, Wright PE. Nat Chem Biol. 2009; 5:789–796. [PubMed: 19841628]
- 11. Wang K, Chodera JD, Yang Y, Shirts MR. J Comp Aided Mol Des. 2013; 27:989-1007.
- Sarich, M.; Prinz, J-H.; Schütte, C. An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation (Advances in Experimental Medicine and Biology. Bowman, GR.; Noé, F.; Pande, VS., editors. Vol. Chapter 3. Springer; 2013. p. 23-44.
- 13. Zuckerman DM. Annu Rev Biophys. 2011; 40:41–62. [PubMed: 21370970]
- 14. Hansmann UH, Okamoto Y. Curr Opin Struct Biol. 1999; 9:177-183. [PubMed: 10322208]
- 15. Pratt LR. J Chem Phys. 1986; 85:5045-5048.
- 16. Dellago C, Bolhuis PG, Chandler D. J Chem Phys. 1998; 108:9236–9245.
- 17. Dellago C, Bolhuis PG, Csajka FS, Chandler D. J Chem Phys. 1998; 108:1964–1977.
- 18. Zuckerman DM, Woolf TB. J Chem Phys. 1999; 111:9475–9484.
- 19. E W, Ren W, Vanden-Eijnden E. Phys Rev B. 2002; 66:052301.
- 20. Laio A, Parrinello M. Proc Natl Acad Sci U S A. 2002; 99:12562-12566. [PubMed: 12271136]
- 21. van Erp TS, Moroni D, Bolhuis PG. J Chem Phys. 2003; 118:7762-7774.
- 22. Faradjian AK, Elber R. J Chem Phys. 2004; 120:10880-10889. [PubMed: 15268118]
- Zhang BW, Jasnow D, Zuckerman DM. Proc Natl Acad Sci U S A. 2007; 104:18043–18048. [PubMed: 17984047]
- 24. Pan AC, Sezer D, Roux B. J Phys Chem B. 2008; 112:3432–3440. [PubMed: 18290641]
- 25. Hawk AT, Konda SSM, Makarov DE. J Chem Phys. 2013; 139:064101. [PubMed: 23947837]
- Hamelberg D, Mongan J, McCammon JA. J Chem Phys. 2004; 120:11919–11929. [PubMed: 15268227]
- Arrar M, de Oliveira CAF, Fajer M, Sinko W, McCammon JA. J Chem Theory Comput. 2013; 9:18–23. [PubMed: 23316122]
- Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. J Comput Chem. 1992; 13:1011–1021.
- 29. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. J Comput Chem. 1995; 16:1339–1350.
- 30. Ferrenberg A, Swendsen R. Phys Rev Lett. 1989; 63:1195–1198. [PubMed: 10040500]
- Gallicchio E, Andrec M, Felts AK, Levy RM. J Phys Chem B. 2005; 109:6722–6731. [PubMed: 16851756]
- 32. Bartels C, Karplus M. J Com Chem. 1997; 18:1450-1462.
- 33. Shirts MR, Chodera JD. J Chem Phys. 2008; 129:124105. [PubMed: 19045004]
- 34. Tan Z. J Am Stat Assoc. 2004; 99:1027-1036.
- Tan Z, Gallicchio E, Lapelosa M, Levy RM. J Chem Phys. 2012; 136:144102. [PubMed: 22502496]
- 36. Zhu F, Hummer G. J Comput Chem. 2012; 33:453-465. [PubMed: 22109354]
- 37. Sugita Y, Okamoto Y. Chem Phys Lett. 1999; 314:141-151.
- 38. Sugita Y, Okamoto Y. Chem Phys Lett. 2000; 329:261-270.
- 39. Sugita Y, Kitao A, Okamoto Y. J Chem Phys. 2000; 113:6042-6051.
- 40. Mitsutake A, Sugita Y, Okamoto Y. Biopolymers. 2001; 60:96–123. [PubMed: 11455545]
- 41. Kokubo H, Tanaka T, Okamoto Y. J Comput Chem. 2011; 32:2810–2821. [PubMed: 21710634]
- 42. Kokubo H, Tanaka T, Okamoto Y. J Comput Chem. 2013; 34:2601–2614. [PubMed: 24006253]

- 43. Mitsutake A, Okamoto Y. J Chem Phys. 2004; 121:2491–2504. [PubMed: 15281846]
- 44. Marinari E, Parisi G. Europhys Letters. 1992; 19:451–458.
- 45. Geyer CJ, Thompson EA. J Am Stat Assoc. 1995; 90:909–920.
- 46. Li H, Fajer M, Yang W. J Chem Phys. 2007; 126:024106. [PubMed: 17228942]
- 47. Huang X, Bowman GR, Pande VS. J Chem Phys. 2008; 128:205106. [PubMed: 18513049]
- 48. Swendsen R, Wang J-S. Phys Rev Lett. 1986; 57:2607-2609. [PubMed: 10033814]
- 49. Hansmann UH. Chem Phys Lett. 1997; 281:140–150.
- Jiang W, Hodoscek M, Roux B. J Chem Theory Comput. 2009; 5:2583–2588. [PubMed: 21857812]
- 51. Jiang W, Roux B. J Chem Theory Comput. 2010; 6:2559–2565. [PubMed: 21857813]
- 52. Jiang W, Luo Y, Maragliano L, Roux B. J Chem Theory Comput. 2012; 8:4672-4680.
- 53. Rodinger T, Howell PL, Pomes R. J Chem Theory Comput. 2006; 2:725-731.
- 54. Rauscher S, Neale C, Pomes R. J Chem Theory Comput. 2009; 10:2640–2662.
- 55. Rhee YM, Pande VS. Biophys J. 2003; 84:775–786. [PubMed: 12547762]
- 56. Shirts M, Pande VS. Science. 2000; 290:1903-1904. [PubMed: 17742054]
- 57. Gallicchio E, Levy RM, Parashar M. J Comp Chem. 2008; 29:788-794. [PubMed: 17876761]
- 58. Affentranger R, Tavernelli I, Iorio E. J Chem Theory Comput. 2006; 2:217-228.
- 59. Wang L, Friesner RA, Berne BJ. J Phys Chem B. 2011; 115:9431-9438. [PubMed: 21714551]
- Wang L, Berne BJ, Friesner RA. Proc Natl Acad Sci U S A. 2012; 109:1937–1942. [PubMed: 22308365]
- 61. Liu P, Kim B, Friesner RA, Berne BJ. Proc Natl Acad Sci U S A. 2005; 102:13749–13754. [PubMed: 16172406]
- 62. Lyman E, Ytreberg FM, Zuckerman DM. Phys Rev Lett. 2006; 96:4.
- 63. Liu P, Voth GA. J Chem Phys. 2007; 126
- Bussi G, Gervasio FL, Laio A, Parrinello M. J Am Chem Soc. 2006; 128:13435–13441. [PubMed: 17031956]
- 65. Kannan S, Zacharias M. Proteins: Struct Funct Bioinf. 2007; 66:697-706.
- 66. Fajer M, Hamelberg D, McCammon JA. J Chem Theory Comput. 2008; 4:1565–1569. [PubMed: 19461870]
- 67. Roitberg AE, Okur A, Simmerling C. J Phys Chem B. 2007; 111:2415–2418. [PubMed: 17300191]
- 68. Okur A, Roe DR, Cui GL, Hornak V, Simmerling C. J Chem Theory Comput. 2007; 3:557–568.
- 69. Li HZ, Li GH, Berg BA, Yang W. J Chem Phys. 2006; 125:5.
- 70. Meng YL, Roitberg AE. J Chem Theory Comput. 2010; 6:1401–1412. [PubMed: 20514364]
- 71. Itoh SG, Damjanovic A, Brooks BR. Proteins: Struct Funct Bioinf. 2011; 79:3420-3436.
- 72. Min D, Chen M, Zheng L, Jin Y, Schwartz MA, Sang Q-XA, Yang W. J Phys Chem B. 2011; 115:3924–3935. [PubMed: 21417256]
- 73. Zheng W, Andrec M, Gallicchio E, Levy RM. Proc Natl Acad Sci U S A. 2007; 104:15340–15345. [PubMed: 17878309]
- 74. Hritz J, Oostenbrink C. J Chem Phys. 2007; 127:204104. [PubMed: 18052416]
- 75. Rosta E, Hummer G. J Chem Phys. 2009; 131:165102. [PubMed: 19894977]
- 76. Smith DB, Okur A, Brooks B. Chem Phys Lett. 2012; 545:118–124. [PubMed: 23087450]
- 77. Trebst S, Troyer M, Hansmann UHE. J Chem Phys. 2006; 124:174903. [PubMed: 16689600]
- Li X, O'Brien CP, Collier G, Vellore NA, Wang F, Latour RA, Bruce DA, Stuart SJ. J Chem Phys. 2007; 127:164116. [PubMed: 17979328]
- 79. Zhang C, Ma J. J Chem Phys. 2008; 129:134112. [PubMed: 19045083]
- 80. Zhang C, Ma J. J Chem Phys. 2009; 130:194112. [PubMed: 19466826]
- 81. Kouza M, Hansmann UHE. J Chem Phys. 2011; 134:044124. [PubMed: 21280704]
- 82. Zhang W, Chen J. J Chem Theory Comput. 2013; 9:2849–2856.
- 83. Zhang W, Chen J. J Comput Chem. 2014; 35:1682–1689. [PubMed: 24995857]
- 84. Sindhikara D, Meng Y, Roitberg AE. J Chem Phys. 2008; 128:024103. [PubMed: 18205439]

- 85. Sindhikara DJ, Emerson DJ, Roitberg AE. J Chem Theory Comput. 2010; 6:2804–2808.
- Plattner N, Doll JD, Dupuis P, Wang H, Liu Y, Gubernatis JE. J Chem Phys. 2011; 135:134111. [PubMed: 21992286]
- 87. Lu J, Vanden-Eijnden E. J Chem Phys. 2013; 138:084105. [PubMed: 23464138]
- 88. Chodera JD, Shirts MR. J Chem Phys. 2011; 135:194110. [PubMed: 22112069]
- 89. Zhang W, Wu C, Duan Y. J Chem Phys. 2005; 123:154105. [PubMed: 16252940]
- 90. Periole X, Mark AE. J Chem Phys. 2007; 126:014903. [PubMed: 17212515]
- 91. Nymeyer H. J Chem Theory Comput. 2008; 4:626–636.
- 92. Abraham MJ, Gready JE. J Chem Theory Comput. 2008; 4:1119–1128.
- 93. Rosta E, Hummer G. J Chem Phys. 2010; 132:034102. [PubMed: 20095723]
- 94. Denschlag R, Lingenheil M, Tavan P. Chem Phys Lett. 2008; 458:244-248.
- 95. Zuckerman DM, Lyman E. J Chem Theory Comput. 2006; 2:1200–1202.
- 96. Plattner N, Doll JD, Meuwly M. J Chem Theory Comput. 2013; 9:4215-4224.
- Zheng W, Andrec M, Gallicchio E, Levy RM. J Phys Chem B. 2008; 112:6083–6093. [PubMed: 18251533]
- 98. Radak B, Romanus M, Gallicchio E, Lee T-S, Weidner O, Deng N-J, He P, Dai W, York D, Levy RM, Jha S. A framework for flexible and scalable replicaexchange on production distributed CI. Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery. 2013:26.
- 99. Gallicchio E, et al. manuscript in preparation.
- 100. Gallicchio E, Lapelosa M, Levy RM. J Chem Theory Comput. 2010; 6:2961–2977. [PubMed: 21116484]
- 101. Gallicchio E, Levy RM. J Comput Aided Mol Des. 2012; 26:505-516. [PubMed: 22354755]
- Lapelosa M, Gallicchio E, Levy RM. J Chem Theory Comput. 2012; 8:47–60. [PubMed: 22368530]
- 103. Wickstrom L, He P, Gallicchio E, Levy RM. J Chem Theory Comput. 2013; 9:3136–3150. [PubMed: 25147485]
- 104. Gallicchio E, Deng N, He P, Wickstrom L, Perryman AL, Santiago DN, Forli S, Olson AJ, Levy RM. J Comput Aided Mol Des. 2014; 28:475–490. [PubMed: 24504704]
- 105. Gilson MK, Given JA, Bush BL, McCammon JA. Biophys J. 1997; 72:1047–1069. [PubMed: 9138555]
- 106. Fukunishi H, Watanabe O, Takada S. J Chem Phys. 2002; 116:9058–9067.
- 107. Gallicchio E, Levy R. J Comput Chem. 2004; 25:479-499. [PubMed: 14735568]
- 108. Gallicchio E, Paris K, Levy RM. J Chem Theory Comput. 2009; 5:2544–2564. [PubMed: 20419084]
- 109. Banks J, et al. J Comp Chem. 2005; 26:1752–1780. [PubMed: 16211539]
- Allen, MP.; Tildesley, DJ. Computer Simulation of Liquids. Oxford University Press; New York: 1993.
- 111. Kullback S, Leibler RA. Ann Math Stat. 1951; 22:79-86.
- 112. Jorgensen WL, Maxwell DS, Tirado-Rives J. J Am Chem Soc. 1996; 118:11225–11236.
- Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. J Phys Chem B. 2001; 105:6474– 6487.
- 114. Zhang BW, et al. manuscript in preparation.
- 115. Bergonzo C, Henriksen NM, Roe DR, Swails JM, Roitberg AE, Cheatham TE 3rd. J Chem Theory Comput. 2014; 10:492–499. [PubMed: 24453949]



Figure 1.

Representation of (λ, T) 2D RE approach to the calculation of binding free energies. Each cell represent an alchemical thermodynamic state at each temperature. The red dashed line illustrates a possible thermodynamic path connecting the bound and unbound conformations of the complex.





(a) Side and top view of b-cyclodextrin-heptonoate complex, (b) binding energy distributions at $\lambda = 1.0$ for T=300 and 200K from two 72ns 1D Sync REMD simulations.



Figure 3.

Simulation results at 300K from 1D Async REMD using different combinations of the MD period and the number of exchanges attempted: (a) binding free energies calculated from UWHAM reweighting; (b) binding energy distributions including all data points from 0 to 40 ns; (c) KL divergences of binding energy distributions to the standard 1D Sync REMD result; (d), (e) and (f) statistical inefficiencies.



Figure 4.

Simulation results from 2D Async REMD using different combinations of the MD period and the number of exchanges attempted: (a) binding energy distributions at 300K; (b) KL divergences at 300K; (c) statistical inefficiency as the function of the number of exchanges attempted; (d) statistical inefficiency as the function of the MD period when the ratio of EX/MD is fixed; (e) binding energy distributions at 200K; (f) KL divergences at 200K.

Xia et al.



Figure 5.

2D Async REMD results from the BOINC distributed network at Temple University: (a) distribution of wall clock times for individual 100ps MD simulations; (b) distribution of the number of exchanges attempted per MD period of 100 ps across 240 replicas; (c) binding energy distributions at 300K; (d) statistical inefficiency.



Figure 6.

Schematic diagram of the asynchronous RE algorithm implemented in the ASyncRE software. The filesystem resides on a coordination server, each cell represent a replica which can be either in a waiting ("W") state or running ("R") state. Replicas in the waiting state can exchange thermodynamic parameters as illustrated by the curved arrows at the bottom of the diagram. Replica in the running state are submitted to the job manager for execution on remote compute resources.

Table 1

Wall Time Information of Async REMD Implemented with BOINC

wall clock time (median)	b-cyclodextrin	b-cyclodextrin	ABL kinase
system size	144 + 22 atoms	144 + 22 atoms	4421+69 atoms
MD period	10 ps	100 ps	10 ps
MD wall time	4m 15s	41m 19s	42m 42s
exchange wall time	4m 39s	14m 35s	16m 26s
overhead	5m 16s	8m 50s	12m 08s
fraction of overhead	37%	13%	17%