

JCTC

Journal of Chemical Theory and Computation

A Second Look at Canonical Sampling of Biomolecules Using Replica Exchange Simulation

Daniel M. Zuckerman* and Edward Lyman†

Department of Computational Biology, School of Medicine, and Department of Environmental & Occupational Health, Graduate School of Public Health, Suite 3064 BST3, 3501 Fifth Avenue, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Received February 10, 2006

Abstract: The replica exchange approach, also called parallel tempering, is gaining popularity for biomolecular simulation. We ask whether the approach is likely to be efficient compared to standard simulation methods for fixed-temperature equilibrium sampling. To examine the issue, we make a number of straightforward observations on how “fast” high-temperature molecular simulations can be expected to run, as well as on how to characterize efficiency in replica exchange. Although our conclusions remain to be fully established, on the basis of a range of results in the literature and some of our own work with a 50-atom peptide, we are not optimistic for the efficiency of replica exchange for the canonical sampling of biomolecules.

Because of growing interest in temperature-based sampling methods for biomolecules, such as replica exchange^{1–9} (see also ref 10), this letter aims to make some observations and raise some potentially important questions which we have not seen addressed sufficiently in the literature. Mainly, we wish to call attention to limits on the maximum speed-up to be expected from temperature-based methods and also note the need for careful quantification of sampling efficiency. Here, *we are strictly concerned with canonical sampling at a fixed temperature*, and *not* with conformational searching. Because potentially lengthy studies may be necessary to address the issues, we felt it would be useful to bring them to the attention of the broader community. Some of the observations we make below were noted previously in a

careful study of a one-dimensional system by Brown and Head-Gordon.¹¹

We will base our discussion around a generic replica exchange protocol, consisting of $M + 1$ levels spanning from the temperature T_0 , at which canonical sampling is desired, up to T_M . Replica exchange is motivated by the increased rate of barrier crossing possible at higher temperatures. We assume each level is simulated for a time t_{sim} , which implies a total CPU cost $(M + 1) \times t_{\text{sim}}$. In typical explicitly solvated peptide systems, $M \sim 20$, $T_0 \approx 300$ K, and $T_M \sim 450$ K.⁵ For typical M values, the relatively low *maximum* temperature (T_M) values reflect the well-known requirement for configuration-space overlap between neighboring levels of the temperature ladder:^{4,5} if temperature increments (and also T_M) become too large, the necessary exchanges between levels become rare because of low overlap. We note that a new exchange variant introduced by Berne and co-workers permits the use of “cold” solvent and larger temperature gaps,¹² but the issues we raise still apply to the new protocol, especially as larger solutes are considered.

While replica exchange is often thought of as an “enhanced sampling method”, what does that mean? Indeed, what is an appropriate criterion for judging efficiency? As our first observation, we believe that (**Obs. I**) efficiency can only mean a decrease in the total CPU usage—that is, *summed over all processors*—for a given degree of sampling quality. (We will defer the necessary discussion of assessing sampling quality and only assume such assessment is possible.) When the goal is canonical sampling at T_0 , after all, one has the option of running an ordinary parallel simulation at T_0 (e.g., ref 13) or perhaps $M + 1$ independent simulations.¹⁴ A truly efficient method must be a superior alternative to such “brute force” simulation.

Reports in the literature offer an ambiguous picture as to whether replica exchange attains efficiency for the canonical sampling of biomolecules (**Obs. II**). Sanbonmatsu and Garcia compared replica exchange to an equivalent amount of brute-force sampling, but their claim of efficiency is largely based on the alternative goal of enhancing sampling over the full range of temperatures, rather than for canonical sampling at T_0 .⁵ When the data solely for T_0 are examined, there is no clear gain, especially noting that assessment was based on principal components derived only from the replica exchange data. Another claim of efficiency, by Duan and co-workers,⁹ fails to include the full CPU cost of all $M + 1$ levels. When suitably corrected, there does appear to be a speed-up of perhaps a factor of 2 for $T_0 = 308$ K, but the system studied is considerably smaller (permitting larger temperature jumps) than would be possible in protein systems of interest. Another efficiency claim by Roe et al. also does not account for the

* Corresponding author phone: 412-648-3335; fax: 412-648-3163; e-mail: dmz@ccb.pitt.edu

† E-mail: elyman@ccb.pitt.edu (E.L.).

Table 1. High-Temperature Speed-up Factors Calculated Using Arrhenius Factors^a

	$\Delta E = 2k_B T_0$	$4k_B T_0$	$6k_B T_0$	$8k_B T_0$
$T_M = 400$ K	1.65	2.72	4.48	7.39
500 K	2.23	4.95	11.0	24.5
600 K	2.72	7.39	20.1	54.6

^a Speed-up factors are computed as the ratio $k_a(T_M)/k_a(T_0 = 300$ K) for the indicated energy barriers ΔE via eq 1. Energy barriers are given in units of $k_B T_0$. A rough estimate of the efficiency factor (the factor by which the total CPU usage is reduced) obtainable in an M -level parallel replica exchange simulation with maximum temperature T_M is the table entry divided by $M + 1$.

full CPU cost of all ladder levels.⁸ In a structural-glass system, replica exchange was found not to be helpful,¹⁵ although efficiency has been noted in spin systems.^{1,16} We emphasize that *biomolecular* replica exchange should indeed be efficient in certain cases (with high enough energy barriers, see below). At least one such instance has been noted by Garcia, using a suitable brute-force comparison system.¹⁷

The lack of clear-cut results in a much-heralded approach merits closer examination. What might be preventing efficiency gain? Or put another way, what is the maximum efficiency possible in a standard replica exchange simulation? The very construction of the method implies that, (**Obs. III**) in any parallel exchange protocol, the sampling “speed” at the bottom level—lowest T —will be controlled by the speed at which the top level—highest T —samples configuration space. A parallel exchange simulation can go no faster than its fastest level and, on average, will be slower. Therefore, given our interest in efficient canonical sampling at T_0 , the speed of the top level should exceed that of the bottom by *at least* a factor of $M + 1$. If not, the simulation does not “break even” in total CPU cost, as compared to brute-force canonical sampling at T_0 for the full length $(M + 1) \times t_{\text{sim}}$.

The basic temperature dependence of barrier-crossing rates is well-known (e.g., ref 18) and has important consequences for replica exchange. The Arrhenius factor indicates that the temperature-dependent rate k for crossing a particular barrier obeys

$$k_a(T) = k_0 \exp(\Delta S/k_B) \exp(-\Delta E/k_B T) \quad (1)$$

for a fixed-volume system, where k_0 is an unknown prefactor insensitive to temperature and is assumed constant; ΔE is the energy barrier, and ΔS is the entropy barrier—that is, “narrowing” of the configuration space—which must be expected in a multidimensional molecular system. Two observations are immediate: (**Obs. IV**) the entropic component of the rate is completely unaffected by an increase in temperature, and the possible speed-up due to the energetic part can easily be calculated.

Table 1 gives possible speed-ups for several energy barriers and temperatures, in units of $k_B T_0$ for $T_0 = 300$ K. Speed-ups are computed simply as the ratio $k_a(T_M)/k_a(T_0)$ for possible values of T_M . It is clear that, for modest barriers, the speed-up attainable even with a top temperature $T_M = 500$ K is only on the order of a typical number of replicas in replica exchange, $M \sim 20$. Thus, (**Obs. V**) if modest barriers ($< 8k_B T_0$) dominate a system’s dynamics, efficiency will be difficult to obtain via a typical replica exchange

simulation, because the speed-up factor noted in the table needs to be divided by $M + 1$.

We reiterate that our discussion is narrowly focused on equilibrium sampling. When, for instance, thermodynamic information over a range of temperatures is desired (e.g., ref 6), replica exchange may indeed be useful. Certainly, if temperatures lower than our T_0 (300 K) are studied, the speed-up factors given in the table will increase.

How high are barriers encountered in molecular systems? We can only begin to answer this question, but one must first be careful about which barriers matter. We believe that (**Obs. VI**) “local” barriers will matter the most; that is, the energy barriers actually encountered along a trajectory will dominate the sampling speed. Apparent barriers determined by projections onto arbitrary low-dimensional reaction coordinates would seem of uncertain value.¹⁹ (We note that Zwanzig has attempted to account for local roughness with an effective diffusion constant on a slowly varying landscape.²⁰)

Evidence from simulations and experiments is far from complete but indicates that (**Obs. VII**) energy barriers in molecular systems appear to be modest. Here, unless noted otherwise, $T_0 \approx 300$ K. In their extensive study of a tetrapeptide, Czerminski and Elber found barriers < 3 kcal/mol $\approx 5k_B T_0$ for the lowest-energy transition path.²¹ Equally interesting, they found approximately 1000 additional paths with similar energy profiles (differing by < 1 kcal/mol $< 2k_B T_0$)—suggesting what we might term a “pebbly” rather than “mountainous” energy landscape (see also ref 22). In our own work (unpublished) with the implicitly solvated 50-atom dileucine peptide,²³ increasing the temperature from 298 K to a series of values (400, 500, and 600 K), independently, led to hopping-rate increases by factors of (2.0, 3.8, and 4.6), respectively. The comparison of these speed-up factors—for hopping between the states defined in ref 24—with the table suggests that there is no barrier larger than $\sim 3k_B T_0$. Similarly, Sanbonmatsu and Garcia found that barriers for explicitly solvated met-enkephalin were small, on the order of $k_B T_0$.⁵ An experimental study has also suggested that barriers are modest ($< 6k_B T_0$).²⁵ Although this list is fairly compelling, we believe the question of barrier heights is far from settled. Further study should carefully consider local versus global barriers, as well as entropy versus energy components of barriers. (We purposely do not discuss barriers to protein folding, because our scope here is solely equilibrium fluctuations.)

Importantly, the goal of understanding efficiency implies the need for reliable means for assessing sampling. An ideal approach to assessment would survey all pertinent substates to ensure appropriate Boltzmann frequencies. Present approaches to assessment typically calculate one- or two-dimensional free energy surfaces (equivalently, projected probability distributions), which are evaluated visually. Principal components (e.g., refs 5 and 9) as well as “composite” coordinates such as the radius of gyration⁸ are popular coordinate choices. Yet we believe that (**Obs. VIII**) the use of low-dimensional sampling assessment is intrinsically limited, because it could readily mask structural diversity—that is, be consistent with substantially distinct conformational ensembles. Future work could usefully pursue higher-dimensional measures, which can always be numerically compared between independent simulations for sam-

pling assessment. In our own work, for instance, we have begun to use a histogram measure which directly reports on the structural distribution of an ensemble.²⁶

Finally, the possibility for improving replica exchange should be noted. A number of groups have made efforts to optimize the approach by examining temperature ladders and exchange acceptance ratios.^{27–30} Such efforts are important and may benefit from the discussion presented here. The work of Brown and Head-Gordon¹¹ and our perspective as reflected in the table suggest that a reduction of the number of replicas M could prove useful (for a fixed T_M). However, the extent to which M can be reduced will depend on a concomitant increase in the exchange attempt frequency, whose cost is hardware-dependent and less amenable to analysis. This point, interestingly, was raised in the 1990 J-walking paper.¹⁰

In conclusion, we have attempted to tie together a number of straightforward observations which reflect concerns about the effectiveness of the replica exchange simulation method, when the goal is single-temperature canonical sampling for biomolecular systems. On the basis of typical current implementations of the replica exchange approach, it is far from clear that the approach is one that should be widely adopted for canonical sampling. The concerns we have outlined suggest that other simulation strategies, such as Hamiltonian exchange³¹ and resolution exchange,^{24,32} may merit consideration—as well as scrutiny. Alternative temperature-based schemes (e.g., refs 10, 11, 12, and 33) also should be considered. We emphasize that our goal has been to raise questions more than to answer them.

ACKNOWLEDGMENT.

The authors wish to thank Rob Coalson, Juan de Pablo, Ron Elber, Angel García, Hagai Meirovitch, Rohit Pappu, and Robert Swendsen for very useful conversations. We gratefully acknowledge support from the NIH, through Grants ES007318 and GM070987. We also greatly appreciate support from the Department of Computational Biology and the Department of Environmental & Occupational Health.

REFERENCES

- (1) Swendsen, R. H.; Wang, J.-S. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609.
- (2) Geyer, C. J. Markov Chain Monte Carlo Maximum Likelihood. In *Proceedings of the 23rd Symposium on the Interface*; Keramidas, E.

- M., Ed.; Computing Science and Statistics Interface Foundation of North America: Fairfax Station, VA, 1991.
- (3) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608.
- (4) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140–150.
- (5) Sanbonmatsu, K. Y.; Garcia, A. E. *Proteins* **2002**, *46*, 225–234.
- (6) Paschek, D.; Garcia, A. E. *Phys. Rev. Lett.* **2004**, *93*, 238105.
- (7) Rathore, N.; Chopra, M.; de Pablo, J. J. *J. Chem. Phys.* **2005**, *122*, 024111.
- (8) Roe, D. R.; Hornak, V.; Simmerling, C. J. *Mol. Biol.* **2005**, *352*, 370–381.
- (9) Zhang, W.; Wu, C.; Duan, Y. *J. Chem. Phys.* **2005**, *123*, 154105.
- (10) Frantz, D. D.; Freeman, D. L.; Doll, J. D. *J. Chem. Phys.* **1990**, *93*, 2769–2784.
- (11) Brown, S.; Head-Gordon, T. *J. Comput. Chem.* **2002**, *24*, 68–76.
- (12) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749–13754.
- (13) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283–312.
- (14) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649–666.
- (15) De Michele, C.; Sciortino, F. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2002**, *65*, 051202.
- (16) Wang, J.-S.; Swendsen, R. *Prog. Theor. Phys. Suppl.* **2005**, *157*, 317–323.
- (17) Garcia, A. E. 2005. Personal communication.
- (18) Atkins, P.; de Paula, J. *Physical Chemistry*, 7th ed.; Freeman: New York, 2002.
- (19) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
- (20) Zwanzig, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2029–2030.
- (21) Czerminski, R.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 5580–5601.
- (22) Hyeon, C.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10249–10253.
- (23) All-atom dileucine peptide (ACE-Leu₂-NME) simulations were performed using TINKER version 4.2, running Langevin Dynamics (91 ps⁻¹ friction coefficient, 1.0 fs time step), with the OPLS-AA force field and GBSA implicit solvent.
- (24) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.
- (25) Nevo, R.; Brumfeld, V.; Kapon, R.; Hinterdorfer, P.; Reich, Z. *EMBO Rep.* **2005**, *6*, 482–486.
- (26) Lyman, E.; Zuckerman, D. M. *Biophys. J.* **2006**, In press. Archived version: <http://www.arXiv.org/abs/physics/0601104> (accessed June 2, 2006).
- (27) Earl, D. J.; Deem, M. W. *J. Phys. Chem. B* **2005**, *108*, 6844–6849.
- (28) Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.
- (29) Kone, A.; Kofke, D. *J. Chem. Phys.* **2005**, *122*, 206101.
- (30) Trebst, S.; Troyer, M.; Hansmann, U. H. E. *J. Chem. Phys.* **2006**, *124*, 174903.
- (31) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042–6051.
- (32) Lyman, E.; Zuckerman, D. M. *J. Chem. Theory Comput.* **2006**, In press.
- (33) Neal, R. M. *Stat. Comput.* **2001**, *11*, 125–139.

CT0600464