JCTC Journal of Chemical Theory and Computation

Simulated Solute Tempering

Robert Denschlag, Martin Lingenheil, Paul Tavan, and Gerald Mathias*

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität, Oettingenstrasse 67, 80538 München, Germany

Received May 26, 2009

Abstract: For the enhanced conformational sampling in molecular dynamics (MD) simulations, we present "simulated solute tempering" (SST) which is an easy to implement variant of simulated tempering. SST extends conventional simulated tempering (CST) by key concepts of "replica exchange with solute tempering" (REST, Liu et al. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749). We have applied SST, CST, and REST to molecular dynamics (MD) simulations of an alanine octapeptide in explicit water. The weight parameters required for CST and SST are determined by two different formulas whose performance is compared. For SST only one of them yields a uniform sampling of the temperature space. Compared to CST and REST, SST provides the highest exchange probabilities between neighboring rungs in the temperature ladder. Concomitantly, SST leads to the fastest diffusion of the simulation system through the temperature space, in particular, if the "even-odd" exchange scheme is employed in SST. As a result, SST exhibits the highest sampling speed of the investigated tempering methods.

Introduction

The generation of equilibrium ensembles for macromolecules by all-atom Monte Carlo (MC) or molecular dynamics (MD) simulations is a challenging task due to the huge computational effort which is generally necessary to guarantee ergodic sampling of relevant observables. The required simulation time depends on the number and on the depths of the local minima in the free energy landscape because here the simulation may get trapped for extended periods of time. If the barriers between the minima are mainly of enthalpic nature, generalized ensemble tempering techniques enable faster barrier crossings and, therefore, alleviate the sampling problem.[1,2]

Two important generalized ensemble tempering algorithms are simulated tempering (ST) and replica exchange (RE).[2] The RE method in its original form[3−5] employs several copies (replicas) of the investigated system at different temperatures $T_0 < T_1 < \cdots < T_{N-1}$. Within this temperature ladder, temperature exchanges between the replicas are periodically attempted. The corresponding probability for an exchange is given by a Metropolis criterion.[6] On the other hand, in simulated tempering,[7,8] only one "replica" diffuses along the temperature ladder, where temperature changes are determined by a Metropolis criterion slightly modified with respect to RE. Although both methods are closely related,[9] RE has attracted much more attention than ST, which is indicated by the large number of RE variants that have been suggested.[10−17] Furthermore, the RE efficiency has been the subject of many studies,[18−23] and numerous RE applications to macromolecules have been presented.[24−33] The main reason for the apparent neglect of ST is that this approach requires estimates for certain *a priori* unknown parameters, the so-called weights, to ensure a uniform sampling of the ensembles at all temperatures. In contrast, RE automatically guarantees a uniform sampling of all temperatures and is, therefore, much simpler to control.

Without prior knowledge on the properties of the simulated system an unbiased tempering algorithm should uniformly cover the chosen temperature space to generate an enhanced statistics at the temperature of interest, which usually is the lowest temperature T_0. Therefore, the average time to shuttle a replica between T_0 and the maximal temperature T_{N-1} should be by orders of magnitude shorter than the simulation time. Addressing this issue, Abraham and Gready recently exam-

* Corresponding author phone: +49-89-2180-9228; fax: +49-89-2180-9202; e-mail: gerald.mathias@physik.uni-muenchen.de. Corresponding author address: LMU München, Lehrstuhl für BioMolekulare Optik, Oettingenstrasse 67, D-80538 München, Germany.

10.1021/ct900274n CCC: $40.75
Published on Web 08/24/2009

ined more than forty published RE simulations for their ability to take full advantage of the tempering.[34] In most cases, they found that the simulation times were too short for a sufficiently frequent shuttling of the replicas along the temperature ladder. The reason for large shuttling times is known: This time usually scales with the square of the number of temperatures,[35] which in turn grows with the square root of the number of degrees of freedom[3] (DOF) in the simulation system. Thus, very long round trip times are expected for simulation systems with many DOF. Typical examples for such systems are macromolecules in explicit solvent. Note that such systems may additionally undergo slow phase transitions, like folding and unfolding, which may drastically increase round trip times.

Various strategies aiming at increasing the shuttling frequency were suggested for RE,[14−16,35−38] which should all be transferable to ST. A first class of strategies targets the optimization of the parameters that characterize the setup of an RE simulation. For example, Sindhikara et al.[38] recently recommended that the times between exchange trials should be as small as possible, which, however, has been disputed.[34] Additionally, rules for optimizing the temperature ladders were suggested.[35,39] A second class of strategies tries to reduce the number of DOF that are relevant for the exchange criterion, for example some strategies switch from an explicit to an implicit solvent description for the exchange trials[16] or by tempering only the areas of interest.[15] With such a reduced number of DOF, the temperature steps within the ladder can be chosen larger and, thus, a given temperature range can be covered by less rungs. As a result, the round trip times are drastically shortened. However, the quoted methods of the second class are not rigorous in terms of statistical physics because here the ensemble is modified in a somewhat arbitrary fashion. Recently, Liu et al.[14] have presented a rigorous strategy called "replica exchange with solute tempering" (REST), which largely eliminates the influence of the solvent DOF on the exchange probabilities by a temperature dependent scaling of the Hamiltonian. In this approach only the Hamiltonian at the target temperature remains unscaled and renders the desired physically meaningful ensemble. This restriction is, however, of minor importance for many applications.

An additional strategy to achieve increased shuttling frequencies is the optimal choice of the tempering method itself. For example, with accurately known weight parameters, ST is more efficient than RE because it provides larger acceptance probabilities on the same temperature ladder.[40−42] At given conditions one may equivalently state that ST requires less rungs in the temperature ladder than RE if both techniques are tuned to the same average acceptance probabilities.

The weights necessary for ST can be estimated by short preparatory simulations bearing the risk, however, that these weights are of insufficient accuracy. Addressing this issue, a recent study by Park and Pande[9] suggests that controlling an ST simulation may be less difficult than previously assumed by Mitsutake and Okamoto.[43] In contrast to the latter authors Park and Pande[9] did not stick to the rough estimates for the weights derived from a set of short preparatory simulations but instead updated the weights during the subsequent ST production run.

Here, inspired by the works of Park and Pande[9] and of Liu et al.,[14] we suggest a variant of ST called simulated solute tempering (SST), which shares key concepts with REST and its sequential variant SREST.[13] As we will demonstrate, SST has the advantages of a most simple implementation and of reducing the required number of rungs within the temperature ladder. Note, that SST should be easily transferrable to hybrid methods which combine different tempering techniques[43] or tempering with other enhanced sampling methods.[44]

We start in "Theory and Methods" with an introduction to replica exchange and its variant REST, which employs a temperature dependent scaling of the Hamiltonian. Then, we present SST along with two procedures for calculating the weights and corresponding update schemes. Subsequently, these techniques are applied to an alanine octapeptide (8ALA) solvated in water. The enthalpic barriers of 8ALA are small, and, hence, the benefit of tempering techniques is limited.[22] However, the small barriers provide fast conformational sampling which makes 8ALA ideally suited to compare different tempering strategies among each other with high statistical accuracy. In particular, we investigate the sampling efficiency and convergence of conventional ST (CST), SST, and REST. In addition, we test two temperature exchange schemes for SST to improve the method further. After the presentation and discussion of the results, we conclude the paper summarizing the key messages.

Theory and Methods

We begin by sketching the replica exchange method and the concept of solute tempering, which will lead us, when combined with simulated tempering, to the SST method.

Replica Exchange. Within both conventional temperature RE[5] (CRE) and REST,[14] N copies (replicas) of the system are simulated at temperatures $T_0 < T_1 < \cdots < T_{N-1}$ and sample the associated canonical ensembles. The set of replicas constitutes a so-called generalized ensemble. After predefined time intervals a temperature exchange between pairs of replicas is tried. A Metropolis criterion[6] determines the exchange probability

$$P_{ij} = \min[1, \exp(\Delta_{ij})] \tag{1}$$

between replicas at T_i and T_j with

$$\Delta_{ij} = \beta_i[E_i(\mathbf{x}_j) - E_i(\mathbf{x}_i)] + \beta_j[E_j(\mathbf{x}_i) - E_j(\mathbf{x}_j)] \tag{2}$$

to preserve the canonical ensembles. Here, $E_k(\mathbf{x}_l)$ is the value of the potential energy function associated with the temperature T_k, which is evaluated at a configuration $\mathbf{x}_l$ resulting from the sampling at T_l; $\beta_k = 1/k_B T_k$ is the inverse temperature where k_B is the Boltzmann constant. For CRE the potential energy is independent of the temperature, and Δ_{ij} reduces to

$$\Delta_{ij} = (\beta_i - \beta_j) \cdot [E(\mathbf{x}_i) - E(\mathbf{x}_j)] \tag{3}$$

In the case of REST, in contrast, the potential energy becomes temperature dependent and has the form

$$E_k(\mathbf{x}) = \lambda_{k,0}E^{pp}(\mathbf{x}) + \lambda_{k,1}E^{ps}(\mathbf{x}) + \lambda_{k,2}E^{ss}(\mathbf{x}) \quad (4)$$

where E^{pp}, E^{ps}, and E^{ss} are the solute−solute, solute−solvent, and solvent−solvent parts of the potential energy function at the target temperature T_0 of the sampling; $\lambda_{k,h}$ are parameters depending on the temperature T_k. We choose

$$\lambda_{k,0} = 1,\ \lambda_{k,1} = \sqrt{\beta_0/\beta_k},\ \lambda_{k,2} = \beta_0/\beta_k \quad (5)$$

where $\lambda_{k,1}$ is the geometric mean of $\lambda_{k,0}$ and $\lambda_{k,2}$ instead of the arithmetic mean $(\beta_0 + \beta_k)/2\beta_k$ originally proposed by Liu et al.[14] With this choice, the required scaling of the electrostatic energy and of the corresponding forces at T_k can be achieved by simply scaling the partial charges of the solvent by a factor $(\beta_0/\beta_k)^{1/2}$. Similar considerations hold for the Lennard-Jones interactions. Thus, this choice of the $\lambda_{k,i}$ is conveniently implemented.

The advantage of the REST approach becomes apparent after a few algebraic operations. Inserting eqs 4 and 5 into eq 2 one finds

$$\begin{aligned}\Delta_{ij} = (\beta_i - \beta_j)[E^{pp}(\mathbf{x}_i) - E^{pp}(\mathbf{x}_j)] \\ + (\sqrt{\beta_0\beta_i} - \sqrt{\beta_0\beta_j})[E^{sp}(\mathbf{x}_i) - E^{sp}(\mathbf{x}_j)]\end{aligned} \quad (6)$$

Thus, the difference Δ_{ij}, which determines the acceptance probability eq 1, is exclusively calculated from the solute−solute and solute−solvent energies, whereas the potential energy E^{ss} of the solvent cancels.

One can quantify the benefit of REST compared with CRE by estimating the number of rungs eliminated from the temperature ladder. Assuming that the solvent DOF do not contribute to the exchange probability at all, the ratio N_{REST}/N_{CRE} of the required rungs is estimated by the lower limit $(n_p/(n_s + n_p))^{1/2}$, where n_s and n_p count the DOF of the solvent and solute, respectively. In the following we discuss how simulated tempering can further reduce the required number of rungs.

Simulated Tempering. In simulated tempering[7,8] (ST), a single system is simulated at a temperature T_i, which belongs to a given temperature ladder $T_0 < \cdots < T_{N-1}$. After given time intervals it is checked whether the system temperature T_i can be changed to T_j, where j is usually $i \pm 1$. For $i \in \{0, N-1\}$ the transition to $j = -1$ or $j = N$ is rejected. For other transitions $i \rightarrow j$, the acceptance probability[41]

$$P_{ij} = \min[1, \exp(\Delta_{ij})] \quad (7)$$

with

$$\Delta_{ij} = [\beta_i E_i(\mathbf{x}) - w_i] - [\beta_j E_j(\mathbf{x}) - w_j] \quad (8)$$

represents a Metropolis criterion similar to that of RE. The weights w_k introduced in eq 8 are commonly set to the configurational parts $\beta_k\tilde{F}_k = -\ln \int \exp[-\beta_k E(\mathbf{x})]d\mathbf{x}$ of the dimensionless free energies $\beta_k F_k$ of the simulation system at the temperatures T_k.[7,8] This choice leads to a uniform sampling of all rungs within a given temperature ladder because, in the ergodic limit, the expected ratio $\rho_k \equiv t_k/t$ of the time t_k spent by the simulation at temperature T_k to the total sampling time t is given by the Boltzmann factor of the generalized ensemble

$$\lim_{t\rightarrow\infty} \rho_k = \frac{\exp[-(\beta_k\tilde{F}_k - w_k)]}{\sum_j \exp[-(\beta_j\tilde{F}_j - w_j)]} \quad (9)$$

leading to $\lim_{t\rightarrow\infty} \rho_k = 1/N$ for $w_j = \beta_j\tilde{F}_j$. Since the w_k are *a priori* unknown, one usually tries to estimate these weights from short preparatory simulations. Note that the w_j can be chosen differently, if a nonuniform sampling is desired.[42]

In conventional simulated tempering (CST), the potential energy function $E_k(\mathbf{x})$ is independent of the temperature T_k, and eq 8 reduces to

$$\Delta_{ij} = (\beta_i - \beta_j)E(\mathbf{x}) - (w_i - w_j) \quad (10)$$

For the new SST method introduced here, we transfer the solute tempering concept of REST to ST and use the energy function given by eqs 4 and 5. Inserting these equations into eq 8 yields

$$\begin{aligned}\Delta_{ij} = (\beta_i - \beta_j)E^{pp}(\mathbf{x}) + (\sqrt{\beta_0\beta_j} - \sqrt{\beta_0\beta_i})E^{ps}(\mathbf{x}) - \\ (w_i - w_j)\end{aligned} \quad (11)$$

Thus, for SST the difference Δ_{ij} is only calculated from the solute−solute and solute−solvent energies, while the potential energy E^{ss} of the solvent cancels. As we will show below the solvent−solvent contributions cancel as well in the computation of the weight differences $w_i - w_j$.

At first glance, ST (CST/SST) seems less attractive than RE (CRE/REST) because ST requires the *a priori* unknown weights w_i. However, ST provides larger average acceptance probabilities than RE[40−42] for a given temperature ladder because RE requires a simultaneous exchange of two replicas, whereas only one replica has to be considered for ST. As a result, ST needs only $1/\sqrt{2}$ times the number of rungs in the temperature ladder than RE to cover a given temperature range with the same acceptance probabilities.[39] Now, we turn to different approaches to determine the required weights w_i.

Determination of the Weights w_k. As we have seen above, a uniform sampling of the various temperatures T_k requires that the weights w_k in eq 8 are the configurational parts $\beta_k\tilde{F}_k$ of the dimensionless free energies, which can be estimated from preparatory simulations. For CST, Park and Pande have presented a formula which yields surprisingly good estimates of the w_k at a negligible computational effort (e.g., by executing a single 10 ps MD simulation at each T_k).[9] These authors replaced the potential energy $E(\mathbf{x})$ in eq 10 by the average potential energy $\langle E\rangle_i$ at T_i to get a "typical" Δ_{ij}^{typ} and demanded that $\Delta_{ij}^{typ} = \Delta_{ji}^{typ}$, which leads to the estimate

$$w_j - w_i \approx (\beta_j - \beta_i)\frac{\langle E\rangle_i + \langle E\rangle_j}{2} \quad (12)$$

and which has been further substantiated by Park.[41] In the Appendix we show an alternative derivation of the "trapezoid rule" eq 12 that is based on the assumption of a constant heat capacity C_V and that $\ln[1 + (T_j - T_i)/T_i]$ can be approximated by $(T_j - T_i)/T_i$.

Replacing the temperature dependent potential energies $E_i(\mathbf{x})$ and $E_j(\mathbf{x})$ in eq 8 by averages $\langle E_i\rangle_i$ and $\langle E_j\rangle_i$ transfers the trapezoid rule of CST to SST. Here, $\langle E_j\rangle_i$ is the energy function E_j evaluated for and averaged over the configurations sampled at T_i. Analogous to CST, also these SST averages yield "typical" differences Δ_{ij}^{typ}. Equation 5 and $\Delta_{ij}^{\text{typ}} = \Delta_{ji}^{\text{typ}}$ lead to

$$w_j - w_i \approx \frac{(\beta_j - \beta_i)(\langle E^{pp}\rangle_i + \langle E^{pp}\rangle_j)}{2} + \frac{(\sqrt{\beta_0\beta_j} - \sqrt{\beta_0\beta_i})(\langle E^{ps}\rangle_i + \langle E^{ps}\rangle_j)}{2} \tag{13}$$

As mentioned in the section "Simulated Tempering", $w_j - w_i$ does not depend on the solvent−solvent interactions E^{ss}. Note that one can choose $w_0 = 0$ because only differences $w_j - w_i$ matter in eq 8.

Whereas the computations of the weights w_i in CST and SST by the trapezoid rules eqs 12 and 13, repectively, are approximate, the relation

$$\exp(-w_i) = \sum_{k=0}^{N-1}\sum_{t=1}^{n_k} \frac{\exp\{-\beta_i \sum_{j=0}^{L-1} \lambda_{i,j} E^j[\mathbf{x}_k(t)]\}}{\sum_{m=0}^{N-1} n_m \exp\{w_m - \beta_m \sum_{j=0}^{L-1} \lambda_{m,j} E^j[\mathbf{x}_k(t)]\}} \tag{14}$$

presented earlier by Kumar et al.[45] for the "weighted histogram analysis method" (WHAM) is "exact" for the already sampled ensemble and provides an unbiased estimator for the true dimensionless free energies.[46] Here, N is again the number of temperatures T_k, n_k is the number of configurations $x_k(1),...,x_k(n_k)$ sampled at T_k, and L counts potential energy terms contributing to the Hamiltonian. For CST there is only one such term and one has $L = 1$, $E^0(\mathbf{x}_k) \equiv E(\mathbf{x}_k)$, and $\lambda_{i,0} = 1$. The Hamiltonian of SST eq 4 distinguishes $L = 3$ energy contributions $E^0(\mathbf{x}_k) \equiv E^{pp}(\mathbf{x}_k)$, $E^1(\mathbf{x}_k) \equiv E^{ps}(\mathbf{x}_k)$, and $E^2(\mathbf{x}_k) \equiv E^{ss}(\mathbf{x}_k)$. With the $\lambda_{i,j}$ chosen as given by eq 5 one finds that E^{ss} cancels in eq 14 like it did in the trapezoid rule eq 13. Equation 14 has to be solved self-consistently and yields successively more and more accurate weights as the statistics is improved by an ongoing sampling. Note that Mitsutake and Okamoto[40,43] previously suggested to compute the free energies required for ST through WHAM equations, which are based on energy histograms. In contrast, eq 14 computes the free energies directly from the sampled energies and, therefore, avoids the errors introduced by the histogram discretization.[45−47]

The WHAM formula eq 14 can be used to identify the errors Δw_i of the trapezoid rules eq 12 for CST and eq 13 for SST. In the limit of ergodic sampling, the errors $\Delta w_i = w_i - w_i^{\text{exact}}$ yield through eq 9 the ratios $\lim_{t\to\infty} \rho_i = \exp(\Delta w_i)/\sum_j \exp(\Delta w_j)$. These ratios will deviate from $1/N$ and, therefore, measure deviations from the desired uniformity of the sampling along the ladder of temperatures T_i. To measure this deviation in our simulations, we introduce the quantities

$$\chi_i \equiv N\frac{t_i}{t} \tag{15}$$

A uniform sampling corresponds to $\chi_i = 1.0$ at all T_i. Thus, the χ_i enable easy comparisons of the sampling uniformity achieved with differently sized temperature ladders. If the errors Δw_i are known, the long time limit of χ_i is

$$\hat{\chi}_i = N\frac{\exp(\Delta w_i)}{\sum_{j=0}^{N-1}\exp(\Delta w_j)} \tag{16}$$

For finite simulations, the χ_i exhibit statistical fluctuations which depend on the number of rungs N in the temperature ladder, on the average exchange probabilities $\bar{P}_{ij}$, and on the total number of exchange trials. For each set of these parameters, one can model an actual ST simulation by a computationally inexpensive MC simulation of a random walk along the N rungs of the temperature ladder. We will use large numbers of such MC simulations to estimate the standard deviations σ_i of the χ_i for the respective simulations. These values show to what extent one may expect convergence of sampling uniformity.

Update Schemes for the Weights w_i. Initial guesses for the weights required in ST can be obtained from short preparatory simulations using the formulas presented in the previous paragraph. The correspondingly limited statistical accuracy of the initial weights may entail a strongly nonuniform sampling along the temperature ladder in subsequent ST production runs. However, one may improve the initial guesses by utilizing the information accumulated in the course of the production run. For this purpose different adaptation schemes were suggested.[48−51]

We used a procedure based on the following considerations: Up to the first update the sampling along the temperature ladder is expected to be far off from uniformity. Consequently, a poor statistics of the potential energy distribution is obtained at some temperature rungs. To avoid an impact of this bad statistics on the estimated weights, we determine the first update by $w_i^{\text{new}} = w_i^{\text{old}} + \ln(t_0/t_i)$ where we set the t_i for rungs that have not been visited at all to a full exchange period. Thus, badly sampled rungs will be preferentially sampled until the next update step. By construction this strategy leads to a uniform sampling but may suffer from slow convergence. Therefore, we subsequently switch to a periodical recomputation of the weights either by the trapezoid rule (eqs 12 and 13) or by the WHAM formula (eq 14). In these recomputations, which are executed after each nanosecond of simulation, we exclusively consider the data from the production run and discarded those of the preparatory simulation.

Exchange Scheme. Having established the determination of the weights steering the exchange probabilities, we now sketch the exchange algorithms employed for ST. A straightforward exchange procedure is to choose randomly between an upward or downward exchange trial with probabilities of 50%. We call this exchange procedure "stochastic even/odd" (SEO) scheme because the corresponding exchange scheme for replica exchange is a stochastic instead of a deterministic choice between two groups of replica pairs.[52] The first group contains all "even" pairs (T_{2n},T_{2n+1}) and the second one all "odd" pairs (T_{2n-1},T_{2n}).

Table 1. Overview of Simulations Conducted[a]

label	trajectory time span/ns	temperature range/K	no. of rungs	solvent scaling	determination of weights	exchange scheme
CRE	18 × 0.1	300−500	18	no	-	DEO
CST	27	300−500	18	no	trapezoid	SEO
REST/A	4 × 0.1	300−500	4	yes	-	DEO
REST/B	5 × 0.1	300−500	5	yes	-	DEO
REST/C	4 × 12	300−500	4	yes	-	DEO
SST/A	27	300−500	4	yes	trapezoid	SEO
SST/B	27	300−500	4	yes	WHAM	SEO
SST/C	12	300−500	4	yes	WHAM	DEO
SST/D	12	300−500	5	yes	WHAM	DEO

[a] The weights for the ST simulations have been determined either by the trapezoid rule, eq 12 for CST and eq 13 for SST, respectively, or by the WHAM formula eq 14.

However, the standard exchange scheme used in replica exchange simulations is characterized by alternate exchange trials between these two groups of replica pairs,[3,34,52,53] which we have called the "deterministic even/odd" (DEO) scheme.[52] Formally, one can express the DEO exchange scheme by a relation which combines the involved temperature indices i and i' at the exchange attempt step s by $i' = i + (-1)^{i+s}$. In the framework of simulated tempering DEO alternately tries to shift the single replica at T_i to T_{i+1} (upward) or T_{i-1} (downward). In the case of a successful exchange, however, the previous exchange direction (upward or downward) is maintained for the next exchange trial and so forth until a temperature exchange fails.

We apply both exchange schemes to investigate their influence on the diffusion of the system through the temperature space. This diffusion can be measured in terms of the average round trip time τ required to travel from T_0 to T_{N-1} and back to T_0. For the SEO scheme applied to a temperature ladder with uniform average acceptance probabilities $\bar{P}_{ij} = \bar{P}$, τ in units of the time between exchange trials is related to the average acceptance probability $\bar{P}$ by[35,52,54,55]

$$\tau_{\mathrm{SEO}} = 2N(N-1)/\bar{P} \tag{17}$$

For DEO, τ is given by

$$\tau_{\mathrm{DEO}} = \tau_{\mathrm{SEO}}(1-\bar{P}) \tag{18}$$

in the limit of large N.[52] Assuming this limit, DEO always provides shorter round trip times than SEO or, equivalently, higher round trip rates τ^{-1}. Thus, the question is to what extent this expectation is confirmed for finite ladder sizes N.

Simulation System and Force Field. We have used MD simulations of a poly alanine octapeptide (8ALA), saturated with an acetyl group at the N-terminus and an *N*-methyl group at the C-terminus, to investigate the benefits of the various algorithms introduced above. The peptide was described by the CHARMM22 force field[56] and solvated in a periodic orthorhombic dodecahedron of 18 Å inscription radius containing 1112 water molecules. For the water molecules we employed the transferable three-point intermolecular potential (TIP3P),[57] modified as suggested by MacKerell et al.[56] for usage with CHARMM22. The initial 8ALA structure was generated using the Molden software[58] by setting the backbone dihedral angles to the values $\phi = -58°$ and $\psi = -47°$ to form an ideal α-helix.

MD Simulation Techniques. The software package EGO-MMVI[59] was used for all MD simulations. The electrostatic interactions were treated combining structure-adapted multipole expansions[60] with a moving-boundary reaction-field approach.[59] Here, the cutoff radius for the explicit evaluation of the electrostatic interactions was 18 Å. Beyond this radius, a dielectric continuum was assumed with a static dielectric constant $\varepsilon_s = 80$. The explicit van der Waals interactions were calculated up to a distance of 10 Å and at larger distances a mean-field approach was applied.[61] A multiple-time-step integration scheme[62] with a fastest time step of 2 fs was used. For bonds that include hydrogen atoms, the corresponding bond lengths were constrained using the M-SHAKE algorithm[63] with relative tolerance of 10^{-6}.

System Preparation. Our simulation system was equilibrated for 100 ps with two Berendsen thermostats[64] (coupling times 0.1 ps) separately keeping the solute and the solvent at 300 K. Additionally, a Berendsen barostat[64] (coupling time 1 ps) steered the system to ambient pressure (1 bar). For the subsequent tempering runs we switched from an *NPT* to an *NVT* ensemble.

Simulation Runs. As listed in Table 1 we carried out three short replica exchange simulations [CRE and REST/(A,B)] serving to estimate the initial weights for the extended simulated tempering simulations CST and SST/A-D. For CRE and CST we used the temperature ladder 300 K, 308 K, 317 K, 326 K, 336 K, 346 K, 356 K, 367 K, 378 K, 390 K, 402 K, 415 K, 428 K, 442 K, 456 K, 470 K, 485 K, and 500 K. We found average acceptance probabilities $\bar{P}_{i,i+1}$ between 5% and 14% for CRE and between 21% and 31% for CST. For REST/(A,C) and SST/A-C we used the ladder 300 K, 350 K, 415 K, and 500 K. The corresponding $\bar{P}_{i,i+1}$ range from 5% to 12% (REST) and from 19% to 28% (SST). For REST/B and SST/D, the five temperatures 300 K, 340 K, 387 K, 440 K, and 500 K were used yielding $\bar{P}_{i,i+1}$ between 15% and 19% (REST) and between 38% and 41% (SST). Exchanges were tried every 0.5 ps in all simulations. Tables S1 and S2 in the Supporting Information provide details about the average exchange probabilities along the various ladders. These values were used for setting up the MC simulations mentioned further above.

The extended simulations REST/C, CST, and SST/(A,B) serve us for comparisons of methods and address, in particular, the applicability of different adaptation schemes

to SST. Furthermore, using the simulations SST/C and SST/D we will study the effects of the chosen exchange scheme (SEO vs DEO) and of the (overall) average exchange probability $\bar{P} = \langle \bar{P}_{ij} \rangle$ on the round trip rates τ^{-1} and the sampling speeds.

Sampling Speed. The main objective of tempering methods is to enhance the sampling speed of the simulation. A corresponding measure for the sampling speed is given by an algorithm recently suggested by Lyman and Zuckerman,[65] which we will denote as LZA. LZA integrates the "volume" in configurational space sampled by a trajectory. The average volume sampled during a given simulation time span provides a measure for the sampling speed.

For 8ALA we define the conformational space by the eight dihedral angles ψ_i spanned by the backbone units $N^i - C^i - C^i_\alpha - N^i$. From a trajectory of the eight-dimensional tuples $(\psi_1,...,\psi_8)$, LZA randomly chooses one tuple and removes it from the trajectory together with all other tuples lying within the sphere of predefined radius r around the chosen tuple. This procedure is repeated until all tuples of the initial trajectory have been removed. The number of steps required is a dimensionless measure for the configurational volume V_c sampled by the trajectory. Because this algorithm is nondeterministic, it is repeated m times, and the corresponding average number n_{lza} of required steps is calculated. For our analyses we choose $r = 25°\sqrt{8}$ and $m = 50$. For a fair comparison between ST and RE, we compute the sampling speed per replica, i.e. one for ST and N for RE.

Results and Discussion

At the start of an ST simulation, initial estimates for the weight parameters w_i are needed. We determined these estimates from preparatory simulations using both the approximate trapezoid rule eq 12 and the asymptotically unbiased WHAM formula eq 14. Now, a first issue is the reliability of the trapezoid rule, which we check using the 100 ps CST simulation.

Reliability of the Trapezoid Rule for CST. Table 2 compares the initial CST weights w_i determined by the trapezoid rule eq 12 from the preparatory CRE simulation with the asymptotically unbiased values calculated by the WHAM formula eq 14. For all T_i the w_i obtained by the two formulas agree quite well. The errors Δw_i of eq 12 never exceed 12% of $k_B T_i$, and, correspondingly, the uniformity measures $\hat{\chi}_i$ are all close to 1.0. Hence, one expects a nearly uniform sampling even if the weights are determined by eq 12. Thus, adaptation schemes, which are based on the trapezoid rule and on the WHAM formula, respectively, should be nearly equivalent for the given system. In the CST simulation we, therefore, applied the trapezoid rule for the periodical recomputation of the w_i.

Representative for the eighteen weights, Figure 1(a) shows the deviation of the weights w_8 and w_{17} from their initial values as a function of the simulation time. The exceptional first update $w_k^{\mathrm{new}} = w_k^{\mathrm{old}} + \ln(t_0/t_k)$, which can be seen as a special case of the update scheme proposed by Zhang and Ma,[51] sizably reduces both weights and reflects the nonuniform sampling within the preceding first nanosecond of the CST simulation. Here, the temperatures $T_8 = 378$ K and $T_{17} = 500$ K apparently have been visited more frequently than $T_0 = 300$ K. The following updates, which rely on eq 12, lead to considerable changes of the weights, which, however, become smaller toward the end of the simulation. After 27 ns the weights seem to be converged within roughly ± 0.1. This is approximately the same magnitude of error as the one introduced by the trapezoid rule.

Figure 1(b) shows the measured (circles) and expected (squares) uniformity measures χ_i and $\hat{\chi}_i$ extracted from the last 25 ns of the CST simulation as functions of the temperature. In contrast, the uniformity data shown in Table

Table 2. Weights Determined from the CRE Simulation[a]

i	T_i	w_i (trapezoid)	w_i (WHAM)	Δw_i	$\hat{\chi}_i$
0	300 K	0.0	0.0	−0.00	1.06
1	308 K	482.58	482.59	−0.01	1.05
2	317 K	991.97	991.99	−0.02	1.04
3	326 K	1468.66	1468.69	−0.03	1.03
4	336 K	1963.47	1963.49	−0.02	1.04
5	346 K	2425.14	2425.17	−0.03	1.03
6	356 K	2856.65	2856.68	−0.03	1.03
7	367 K	3299.53	3299.56	−0.03	1.03
8	378 K	3712.36	3712.41	−0.05	1.00
9	390 K	4131.64	4131.73	−0.09	0.97
10	402 K	4521.49	4521.56	−0.07	0.99
11	415 K	4913.99	4914.07	−0.08	0.98
12	428 K	5278.33	5278.44	−0.11	0.95
13	442 K	5642.42	5642.50	−0.08	0.98
14	456 K	5980.14	5980.21	−0.07	0.99
15	470 K	6294.02	6294.09	−0.07	0.99
16	485 K	6606.46	6606.55	−0.09	0.97
17	500 K	6896.68	6896.80	−0.12	0.94

[a] The weights w_i determined by the trapezoid rule eq 12 and by the WHAM formula eq 14 together with the deviations Δw_i and the correspondingly predicted (cf. eq 16) uniformity measures $\hat{\chi}_i$. The WHAM weights were employed as starting values for the CST simulation.

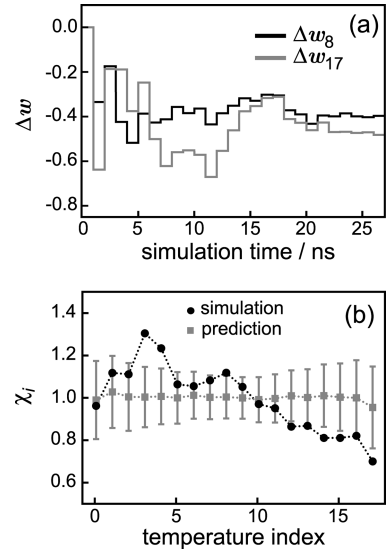

Figure 1. Uniformity of the temperature sampling in the CST simulation. (a) Time evolution of the weights w_8 and w_{17} with respect to their initial values. (b) Uniformity measures observed (χ_i, eq 15, circles) and predicted ($\hat{\chi}_i$, eq 16, squares) after 27 ns at the temperatures T_i. The standard deviations were estimated from MC trial simulations (see text for further details). The dotted lines serve as a guide for the eye.

2 had been extracted from the much shorter preparatory CRE simulation. According to eq 16 the observables $\hat{\chi}_i$ reflect the average deviations Δw_i between the trapezoid weights used during the CST simulation and WHAM weights calculated *a posteriori* from that simulation. As one sees in the figure, the expectation values $\hat{\chi}_i$ are close to one demonstrating that the trapezoid rule induces only small errors into the weights. These data confirm the claim[9] that the trapezoid rule is appropriate for choosing the weights in conventional ST simulations.

Despite the expected nearly uniform sampling of the sampling along the temperature ladder the values χ_i measured for the CST simulation deviate substantially from one, yielding a root-mean-square deviation (RMSD) from uniformity of 16%. Because the weights are calculated with a reasonable accuracy, this nonuniformity of the sampling must be due to a too short CST simulation time. We checked this issue by the simple MC model for the CST simulation described in Theory and Methods, because here the expected deviations from a uniform sampling can be reliably determined.

The gray bars in Figure 1(b) measure the standard deviations σ_i of the χ_i resulting from 1000 MC model simulations covering the same number of exchange trials as our CST simulation. One would now expect that $\mathrm{erf}(\sigma) = 68\%$ of the χ_i are found at smaller deviations than σ_i. In fact, 12 of the 18 χ_i are within the corridor marked by the σ_i, which nicely reproduces the expected statistics. Quite clearly the standard deviations σ_i can be reduced by extending the simulation time, which will then also lead to a CST sampling close to uniformity. Now the question is whether one can estimate the simulation time required for a reasonably uniform sampling. This issue can be addressed by considering the round trip rate τ^{-1} given by eq 17.

CST Round Trip Rates. From the MC model simulations we calculated an average round trip rate τ^{-1} of 0.83 ns^{-1} with a standard deviation of 0.1 ns^{-1}. Equation 17 gives an exact expression for τ applying to the SEO exchange scheme used in the CST simulation. This expression rests on the assumption of identical acceptance probabilities $\bar{P}$ for exchanges along the ladder. The $\bar{P}$ determined from the CST simulation is about 26%, and the resulting value $\tau^{-1} = 0.85$ ns^{-1} is very close to the MC result.

However, the round trip rate observed in the CST simulation is sizably smaller measuring 0.64 ns^{-1}. This deviation suggests that the time interval of 0.5 ps between subsequent exchange trials is too short to yield statistically independent configurations, i.e. that the autocorrelation time of the energy exceeds 0.5 ps. Thus, the system still has some memory of the previous exchange trial, which, however, is tolerable for most practical purposes. Furthermore, a round trip rate of 0.64 ns^{-1} means that only 16 round trips were counted during the CST simulation which is the main cause for the observed 16% RMSD from uniform sampling. To half this RMSD, a 4-fold number of round trips and, thus, a 4-fold simulation time would be necessary. Accordingly, one can estimate the number of round trips needed to achieve a desired level of uniform sampling. In turn, one can *a priori* estimate the required simulation time by multiplying this number by the predicted round trip time given in eq 17.

Table 3. Weights Determined from the REST/A Simulation[a]

i	T_i	w_i (trapezoid)	w_i (WHAM)	Δw_i	$\hat{\chi}_i$
0	300K	0.0	0.0	0.00	1.12
1	350K	−16.91	−16.89	−0.02	1.09
2	415K	−35.64	−35.52	−0.12	0.99
3	500K	−56.16	−55.83	−0.33	0.80

[a] The weights w_i were determined from the initial 100 ps of the REST/A simulation by the trapezoid rule eq 13 and the WHAM formula eq 14 together with the deviations Δw_i and the corresponding uniformity measures $\hat{\chi}_i$. The WHAM weights serve as starting values for the SST/A and SST/B simulation.

Next, we will study the sampling behavior of SST for which we will additionally examine the adaptation scheme based on the WHAM formula.

Reliability of the Trapezoid Rule for SST. Table 3 shows initial weights w_i determined from the short REST/A simulation. Because the solvent−solvent interactions do not contribute to the partition function of SST, these weights are tiny compared to those given in Table 2. Furthermore, the deviations Δw_i between the trapezoid rule eq 13 and the WHAM formula eq 14 are much larger than those listed in Table 2. The associated uniformity measures $\hat{\chi}_i$ predict that errors of this size will lead to a considerable nonuniformity of the SST sampling if the w_i are calculated by the trapezoid rule. Recall here that this rule can be derived based on two assumptions: (i) the heat capacity at consecutive temperatures T_i and T_{i+1} is constant and (ii) the logarithm $\ln(1 + \Delta T_{i+1,i}/T_i)$ is well approximated by its first order Taylor expansion. These conditions are harder to fulfill for SST than for CST because here the temperature steps $\Delta T_{i+1,i}$ are larger.

Figure 2 compares the effects of applying the trapezoid and WHAM rules, respectively, for updating the w_i during SST simulations. Figure 2(a) shows the deviation of the weight w_3 belonging to $T_3 = 500$ K during the SST/A (gray line, trapezoid) and SST/B (black line, WHAM) simulations from the initial value. The first update drastically changes w_3 in both simulations indicating that w_3 has been poorly estimated by the preparatory simulation. The subsequent updates reduce the large initial change to a final deviation of about 1.4 in both cases.

At first glance, the small difference of the resulting w_3 values suggests that the errors of the trapezoid rule are much smaller than predicted by the preparatory REST/A simulation. To check this issue, we have recalculated the weights w_i of the SST/A simulation *a posteriori* by the WHAM formula. The resulting time evolution of w_3 is depicted in Figure 2(a) by the gray dotted line. The difference of 0.22 ± 0.02 between the dotted gray and the solid gray lines is nearly constant during the simulation. Obviously, the trapezoid rule systematically underestimates w_3. A similar underestimate appears already in the initial guess for Δw_3 given in Table 2. Because of this systematic error of the trapezoid rule, the uniformity of the temperature sampling is expected to be suboptimal in SST/A.

Figure 2(b) shows the measured (circles) and predicted (squares) uniformity measures χ_i and $\hat{\chi}_i$ of the SST/A simulation. As indicated by the squares, the average deviations Δw_i between the trapezoid and the WHAM rules predict deviations of up to 13% from uniformity. The measured χ_i

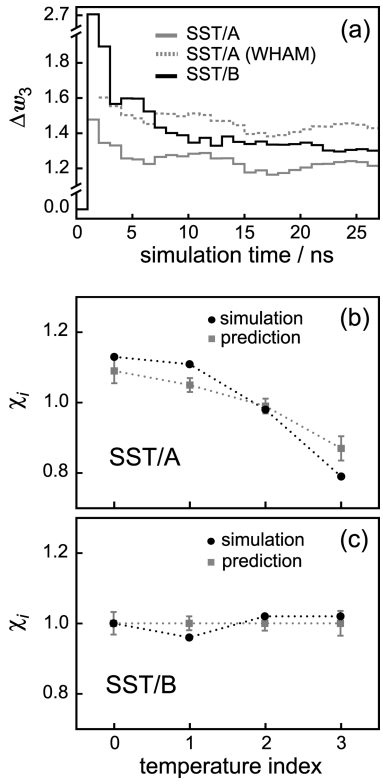

Figure 2. Uniformity of temperature sampling in SST simulations. (a) Deviation of the weight w_3 from its initial value in simulation SST/A (gray) and SST/B (black), respectively. The dotted line shows w_3' calculated a *posteriori* from SST/A using the WHAM expression eq 14. The broken w_3 axis serves to simplify the comparison with Figure 1(a). (b) Measured and predicted uniformity measures χ_i and $\hat{\chi}_i$ of SST/A and (c) of SST/B.

(circles) essentially follow these expectations but show even larger deviations from uniformity, yielding an RMSD of 14%. For instance, χ_3 happens to deviate by about two standard deviations from the respective expectation value $\hat{\chi}_3$. Like for CST, the standard deviations shown as gray bars in the figure were determined from additional MC simulations.

Figure 2(c) compares the uniformity measures of the SST/B simulation. Because the WHAM formula is the reference, the errors Δw_i vanish and eq 16 predicts a uniform sampling $\hat{\chi}_i = 1.0$ at all temperatures. In fact, the measured χ_i are close to 1.0 and show an RSMD of only 3%. Thus, SST/B exhibits an almost perfectly uniform sampling implying that the WHAM formula should be used in SST simulations for updating the w_i. The remaining deviations from uniformity are consistent with the narrow range of the statistical fluctuations estimated by our separate MC simulations. Compared with CST, the much smaller deviations of the χ_i from the predictions $\hat{\chi}_i$ indicate that many more round trips must have occurred during the SST simulations.

SST Round Trip Rates. Equation 17 predicts a round trip rate of 20 ns^{-1} for the simulations SST/(A,B), if the measured average acceptance probability $\bar{P} = 24\%$ is used. Our MC models of SST/(A,B) have reproduced this rate. For the MD simulations SST/A and SST/B, however, we found round trip rates of only 15.8 ns^{-1} and 16.5 ns^{-1}, respectively. Thus, the SST simulations apparently display

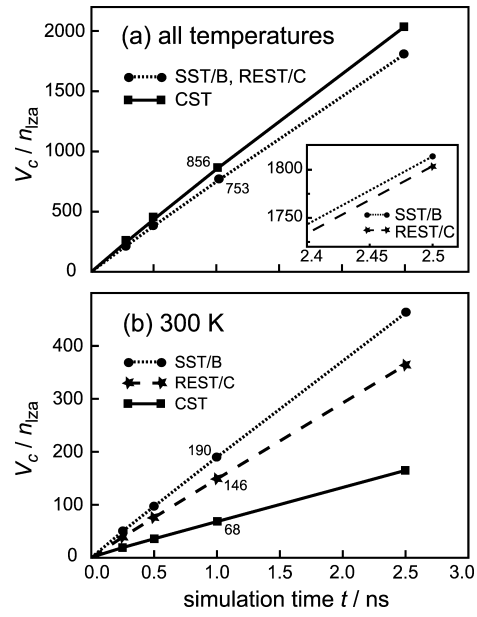

Figure 3. Volumes V_c sampled within 0.25 ns, 0.5 ns, 1 ns, and 2.5 ns by the simulations CST, SST/B, and REST/C. Due to the linearity of the shown $V_c(t)$ curves, the values at 1 ns represent the sampling speeds S in units of n_{lza}/ns. (a) Volumes $V_c(t)$ sampled by the trajectories at all temperatures. According to the inset the $V_c(t)$ curves of REST/C and SST/B are so close that they cannot be distinguished in the main plot. (b) Volume $V_c(t)$ sampled at the target temperature $T_0 =$ 300 K. Note that the statistical errors of the measured volumes are smaller than the symbol sizes.

the same memory effect which was already observed in the CST simulation and which reduces the round trip rates by 20%. Nevertheless, the SST round trip rates are by a factor of 26 larger than the CST rates. Correspondingly, the χ_i are much better converged in SST than in CST.

The large round trip rates and the nearly uniform sampling achieved in simulation SST/B lead to the expectation that this simulation setting leads for a peptide in solution to an improved statistics. We now will address this issue for our sample peptide 8ALA in TIP3P water.

Sampling Speed of CST, SST, and REST. The purpose of any tempering algorithm is to increase the sampling speed within the conformational space of the studied system. For our various simulation settings we determined the sampling speed using the LZA algorithm described in the Theory and Methods section. This iterative algorithm measures the volume $V_c(t)$ of the configuration space sampled within a simulation time t through an average number n_{lza} of iterations. The simulation speed $S(t)$ is then given by the time derivative of $V_c(t)$.

Figure 3(a) shows the volumes V_c sampled by the simulations CST, SST/B, and REST/C at all temperatures as functions of the simulation time t. The respective sampling speeds S are the constants $V_c(t)/t$ at $t = 1$ ns. The V_c curves of REST/C and SST/B cannot be graphically distinguished at the given scale as is documented by the inset in Figure 3(a). Apparently, CST provides the highest overall sampling speed of the three simulations. The sampling speeds of SST/B and REST/C are by about 10% smaller, which may be caused

by the scaling of the solvent part of the Hamiltonian corresponding to an effectively cooler environment.

We have checked the latter conjecture by two MD simulations at 500 K (data not shown) with and without solvent scaling. Here, the effectively cooler solvent indeed reduces the sampling speed of 8ALA by about 20%. For lower temperatures we expect this effect to be correspondingly smaller. However, this small effect is tolerable if the sampling speed at the target temperature $T_0 = 300$ K is sufficiently enhanced, which is after all the aim of solute tempering methods.

Figure 3(b) compares the sampling speeds at $T_0 = 300$ K for the three methods. In contrast to the sampling speed of the generalized ensemble, at 300 K REST/C samples the peptide conformations 2.1 times faster than CST, and SST/B outperforms CST even by a factor of 2.8. Thus, SST/B samples also faster than REST/C, although the two simulations employ the same temperature ladder and the same solvent scaling. An explanation of this speedup is given by the different round trip rates of 11.1 ns^{-1} for REST/C and 16.4 ns^{-1} for SST/B. Due to the higher rate, SST delivers the structural information that is gathered at higher temperatures faster to the target temperature implying an enhanced speed of conformational sampling at T_0.

The reduced round trip rate of REST/C compared to SST/B directly results from the fact that for a given temperature ladder the average acceptance probabilities $\bar{P}_{RE}$ of RE methods (including their sequential versions[13]) are smaller than the probabilities $\bar{P}_{ST}$ of the corresponding ST methods.[40−42] The reason is that in RE the configurations of two replicas must simultaneously meet a certain energy criterion instead of only one replica in ST. Therefore, the average acceptance probability $\bar{P}_{RE}$ should be approximately the square of $\bar{P}_{ST}$. For example, the average acceptance probability of SST/B is about 26%. Thus, we expect a probability of 7% ($0.26^2 \approx 0.07$) for REST/C which is close to the measured value of 9%.

Optimal Exchange Scheme. Because the acceptance probability of REST/C is much smaller than that of SST/B, eq 17 predicts likewise different round trip rates. Compared with that expectation the round trip rate measured for REST/C (11.1 ns^{-1}) seems to be too high compared to SST/B (16.5 ns^{-1}). This large REST/C rate illustrates the advantage of the employed DEO exchange scheme compared to the SEO scheme of SST/B.[52] Furthermore, the optimal exchange probabilities $\bar{P}$ which yield the highest round trip rates are different for these two schemes. For SEO the optimal $\bar{P}$ is 23%,[52,66] whereas for DEO the optimal $\bar{P}$ is between 40% and 45% depending on the ladder size.[39,52]

To investigate the effects of the exchange scheme and the acceptance probability on the SST round trip rates we have carried out the two 12 ns simulations SST/C and SST/D. SST/C switches from SEO to DEO, and SST/D additionally uses five instead of four rungs to span the temperature range from 300 K up to 500 K (see Table 1). SST/D thereby increases $\bar{P}$ to about 40%. For SST/C we found a round trip rate of 20.0 ns^{-1} which is about 20% larger than that of SST/B. Thus, the DEO exchange scheme indeed speeds up the round trips sizably. An additional increase of $\bar{P}$ in simulation SST/D leads to a still larger round trip rate of 23.8 ns^{-1}.

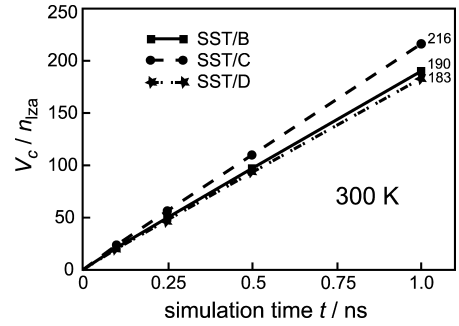

Figure 4. Sampled volumes for the simulations SST/B, SST/C, and SST/D at $T_0 = 300$ K by SST/C. Compared to SST/B the DEO exchange scheme increases the sampling speed by about 15%. This increase corresponds to the enhancement of the round trip rate. Interestingly, the sampling speed of SST/D is smaller than that of SST/B despite the much larger round trip rate. Here, the lower sampling speed is caused by the 25% reduced sampling time at 300 K due to the additional temperature rung, which is not compensated by the higher round trip rate.

Table 4. Round Trip Rates and Sampling Speeds Measured at Temperature $T_0 = 300$ K

label	round trip rates/ns^{-1}	speed/n_{lza}/ns
CST	0.63	68
REST/C	11.1	146
SST/A	15.8	204
SST/B	16.5	190
SST/C	20.0	216
SST/D	23.8	183

Figure 4 shows the effects of the round trip rates, which increase in the sequence SST/B, SST/C, and SST/D, on the sampling speed at 300 K. The highest sampling speed is achieved by SST/C. Compared to SST/B the DEO exchange scheme increases the sampling speed by about 15%. This increase corresponds to the enhancement of the round trip rate. Interestingly, the sampling speed of SST/D is smaller than that of SST/B despite the much larger round trip rate. Here, the lower sampling speed is caused by the 25% reduced sampling time at 300 K due to the additional temperature rung, which is not compensated by the higher round trip rate.

Finally, Table 4 summarizes the round trip rates and sampling speeds measured in our simulations. Using REST instead of CST speeds up the sampling by a factor of 2, whereas SST yields a speedup factor of 3. Note, however, that these factors are conservative estimates because of the particular choice of the 8ALA system. In this system the enthalpic barriers are small, and, therefore, the benefit of tempering methods is limited (see Introduction).[22] Correspondingly, the sampling speeds do not increase very much upon heating the system from the lowest to the highest temperature. However, target applications of tempering methods feature large enthalpic barriers[14,67] for which the sampling has to be accomplished mainly at high temperatures. Correspondingly, the sampling speed at T_0 should depend much stronger on the round trip rates. As a key result,

SST in combination with the DEO exchange scheme should generally show much better sampling properties than REST or CST.

Conclusion

We have introduced simulated solute tempering (SST) which combines the (serial) simulated tempering method with solute tempering, i.e. the key idea of the REST approach. SST poses an efficient alternative to conventional simulated tempering and replica exchange, including REST and its sequential version SREST because it offers the largest acceptance probabilities for a given temperature ladder.

From a practical point of view, it is gratifying to note that SST can be easily implemented. For example, for rigid models of the solvent molecules only the partial charges and the van der Waals parameters have to be scaled to generate the modified Hamiltonians at higher temperatures. Furthermore, SST enables a parallel sampling of many replicas even on heterogeneous computer clusters, because all replicas travel independently through temperature space.

The necessary ingredients of SST are the weights, i.e. the dimensionless free energies of the system at the rungs of the temperature ladder. The trapezoid rule recently suggested by Park and Pande[9] for the computation of the weights is not accurate enough for SST but well suited for CST. Our rederivation of this rule has shown that it is only accurate for temperature ladders featuring small temperature differences, which is the case for CST. In SST a few rungs suffice to span a large temperature range. Due to the failure of the trapezoid rule, the SST weights should be updated using the more complex but asymptotically unbiased WHAM formula of Kumar et al.[45] Then an almost perfectly uniform sampling of the temperature rungs is guaranteed, if the simulation time exceeds the average round trip time by about 2 orders of magnitude.

Our comparison of different sampling methods (REST, CST, SST) applied to an octapeptide in explicit water has demonstrated that the SST sampling is the most efficient one as was shown by the largest round trip rate and the highest sampling speed at T_0. Finally we have shown that the round trip rates can be maximized by using the DEO instead the SEO exchange scheme and by choosing a temperature ladder that provides acceptance probabilities close to 45%.

In conclusion, the sampling efficiency of SST as well as its ease of implementation and application nourishes the hope that simulated tempering will become more popular and that we may see many exciting applications in the future.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft (Grants SFB 533/C1 and SFB 749/C4). Computer time provided by Leibniz Rechenzentrum (project uh408) is gratefully acknowledged.

Appendix

For a canonical ensemble with a heat capacity C_V independent of the temperature, eq 12 can be derived by the following physical considerations. Using the shorthand notation $\Delta X_{ji} \equiv X_j - X_i$, the entropy difference $\Delta S_{ji} = C_V \ln(1 + \Delta T_{ji}/T_i)$ can be estimated by a first order Taylor expansion of the logarithm as $\Delta S_{ji} \approx C_V \Delta T_{ji}/T_i$. With $C_V = \Delta U_{ji}/\Delta T_{ji}$ one gets

$$\Delta S_{ji} \approx \frac{\Delta U_{ji}}{T_i} \tag{19}$$

With the Helmholtz free energy $F = U - TS$, where U denotes the internal energy, the free energy difference ΔF_{ji} between the systems at T_j and T_i can be written as

$$\Delta F_{ji} = \Delta U_{ji} - T_i \Delta S_{ji} - S_i \Delta T_{ji} - \Delta T_{ji} \Delta S_{ji} \tag{20}$$

Inserting eq 19 one immediately finds

$$F_j \approx F_i - S_i \Delta T_{ji} - \Delta U_{ji} \frac{\Delta T_{ji}}{T_i} \tag{21}$$

With eq 21, the dimensionless free energy difference $\Delta\phi_{ji} = F_j/k_B T_j - F_i/k_B T_i$ can be written as

$$\Delta\phi_{ji} \approx \Delta\beta_{ji}(U_i + \Delta U_{ji}) \tag{22}$$

Interchanging i and j one obtains an equally valid estimate

$$\Delta\phi_{ji} \approx \Delta\beta_{ji}(U_j - \Delta U_{ji}) \tag{23}$$

where we have used $\Delta X_{ij} = -\Delta X_{ji}$. An even better approximation is then given by the arithmetic mean

$$\Delta\phi_{ji} \approx \Delta\beta_{ji} \frac{U_i + U_j}{2} \tag{24}$$

where ΔU_{ji} cancels. Restricting the internal energy U to its configurational part, i.e., to the average potential energy $\langle E \rangle$, yields the "trapezoid" rule eq 12.

Supporting Information Available: Average acceptance probabilities $\bar{P}_{i,i\pm1}$ for the various simulations (Tables S1 and S2). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers (Peptide Sci.)* **2001**, *60*, 96.

(2) Okamoto, Y. *J. Mol. Graphics Modell.* **2004**, *22*, 425.

(3) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.

(4) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140.

(5) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.

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

(7) Lyubartsev, A. P.; Martinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.

(8) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451.

(9) Park, S.; Pande, V. S. *Phys. Rev. E* **2007**, *76*, 016703.

(10) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042.

(11) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058.

(12) Affentranger, R.; Tavernelli, I. *J. Chem. Theory Comput.* **2006**, *2*, 217.

(13) Hagen, M.; Kim, B.; Liu, P.; Friesener, R. A.; Berne, B. J. *J. Phys. Chem. B* **2007**, *111*, 1416.

(14) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749.

(15) Kubitzki, M. B.; de Groot, B. L. *Biophys. J.* **2007**, *92*, 4262.

(16) Xu, W.; Lai, T.; Yang, Y.; Mu, Y. *J. Chem. Phys.* **2008**, *128*, 175105.

(17) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.

(18) Rao, F.; Caflisch, A. *J. Chem. Phys.* **2003**, *119*, 4035.

(19) Zhang, W.; Wu, C.; Duan, Y. *J. Chem. Phys.* **2005**, *123*, 154105.

(20) Rick, S. W. *J. Chem. Theory Comput.* **2006**, *2*, 939.

(21) Periole, X.; Mark, A. E. *J. Chem. Phys.* **2007**, *126*, 014903.

(22) Zuckerman, D. M.; Lyman, E. *J. Chem. Theory Comput.* **2006**, *2*, 1200.

(23) Denschlag, R.; Lingenheil, M.; Tavan, P. *Chem. Phys. Lett.* **2008**, *458*, 244.

(24) Zhou, R.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931.

(25) Sanbonmatsu, K. Y.; García, A. E. *Proteins* **2002**, *46*, 225.

(26) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7587.

(27) Cecchini, M.; Rao, F.; Seeber, M.; Caflisch, A. *J. Chem. Phys.* **2004**, *121*, 10748.

(28) Jas, G. S.; Kuczera, K. *Biophys. J.* **2004**, *87*, 3786.

(29) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241.

(30) Nguyen, P. H.; Stock, G.; Mittag, E.; Hu, C.-K.; Li, M. A. *Proteins* **2005**, *61*, 795.

(31) Villa, A.; Stock, G. *J. Chem. Theory Comput.* **2006**, *2*, 1228.

(32) Schrader, T. E.; Schreier, W. J.; Cordes, T.; Koller, F. O.; Babitzki, G.; Denschlag, R.; Renner, C.; Lweneck, M.; Dong, S.-L.; Moroder, L.; Tavan, P.; Zinth, W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15729.

(33) Villa, A.; Widjajakusuma, E.; Stock, G. *J. Phys. Chem.* **2008**, *112*, 134.

(34) Abraham, M. J.; Gready, J. E. *J. Chem. Theory Comput.* **2008**, *4*, 1119.

(35) Nadler, W.; Hansmann, U. H. E. *J. Phys. Chem. B* **2008**, *112*, 10386.

(36) Calvo, F. *J. Chem. Phys.* **2005**, *123*, 124106.

(37) Brenner, P.; Sweet, C. R.; VonHandorf, D.; Izaguirre, J. A. *J. Chem. Phys.* **2007**, *126*, 074103.

(38) Sindhikara, D.; Meng, Y.; Roitberg, A. E. *J. Chem. Phys.* **2008**, *128*, 024103.

(39) Denschlag, R.; Lingenheil, M.; Tavan, P. *Chem. Phys. Lett.* **2009**, *473*, 193.

(40) Mitsutake, A.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *332*, 131.

(41) Park, S. *Phys. Rev. E* **2008**, *77*, 016709.

(42) Zhang, C.; Ma, J. *J. Chem. Phys.* **2008**, *129*, 134112.

(43) Mitsutake, A.; Okamoto, Y. *J. Chem. Phys.* **2004**, *121*, 2491.

(44) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2006**, *128*, 13435.

(45) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.

(46) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.

(47) Kobrak, M. N. *J. Comput. Chem.* **2003**, *24*, 1437.

(48) Berg, B. A. *J. Stat. Phys.* **1996**, *82*, 323.

(49) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450.

(50) Park, S.; Ensign, D. L.; Pande, V. S. *Phys. Rev. E* **2006**, *74*, 066703.

(51) Zhang, C.; Ma, J. *Phys. Rev. E* **2007**, *76*, 036708.

(52) Lingenheil, M.; Denschlag, R.; Mathias, G.; Tavan, P. *Chem. Phys. Lett.* 2009. in press (doi:10.1016/j.cplett.2009.07.039).

(53) Okabe, T.; Kawata, M.; Okamoto, Y.; Mikami, M. *Chem. Phys. Lett.* **2001**, *335*, 435.

(54) Gardiner, C. W. *Handbook of Stochastic Methods*, 2nd ed.; Springer, Berlin, 1985.

(55) Nadler, W.; Hansmann, U. H. E. *Phys. Rev. E* **2007**, *75*, 026109.

(56) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586.

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

(58) Schaftenaar, G.; Noordik, J. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123.

(59) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. *J. Chem. Phys.* **2003**, *118*, 10847.

(60) Niedermeier, C.; Tavan, P. *J. Chem. Phys.* **1994**, *101*, 734.

(61) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: Oxford, 1987.

(62) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. *J. Comput. Chem.* **1997**, *18*, 1729.

(63) Kraeutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Comput. Chem.* **2001**, *22*, 501.

(64) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(65) Lyman, E.; Zuckerman, D. M. *Biophys. J.* **2006**, *91*, 164.

(66) Kone, A.; Kofke, D. A. *J. Chem. Phys.* **2005**, *122*, 206101.

(67) Reichold, R.; Fierz, B.; Kiefhaber, T.; Tavan, P. Submitted for publication.

CT900274N