JCTC Journal of Chemical Theory and Computation

g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates

Jochen S. Hub,*,[†] Bert L. de Groot,[‡] and David van der Spoel[†]

Department of Cell and Molecular Biology, Uppsala University, Box 596, 75124 Uppsala, Sweden, and Computational Biomolecular Dynamics Group, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Abstract: The Weighted Histogram Analysis Method (WHAM) is a standard technique used to compute potentials of mean force (PMFs) from a set of umbrella sampling simulations. Here, we present a new WHAM implementation, termed g_wham, which is distributed freely with the GROMACS molecular simulation suite. g_wham estimates statistical errors using the technique of bootstrap analysis. Three bootstrap methods are supported: (i) bootstrapping new trajectories based on the umbrella histograms, (ii) bootstrapping of complete histograms, and (iii) Bayesian bootstrapping of complete histograms, that is, bootstrapping via the assignment of random weights to the histograms. Because methods ii and iii consider only complete histograms as independent data points, these methods do not require the accurate calculation of autocorrelation times. We demonstrate that, given sufficient sampling, bootstrapping new trajectories allows for an accurate error estimate. In the presence of long autocorrelations, however, (Bayesian) bootstrapping of complete histograms yields a more reliable error estimate, whereas bootstrapping of new trajectories may underestimate the error. In addition, we emphasize that the incorporation of autocorrelations into WHAM reduces the bias from limited sampling, in particular, when computing periodic PMFs in inhomogeneous systems such as solvated lipid membranes or protein channels.

Introduction

The concept of potentials of mean force (PMFs), originally introduced by Kirkwood,[1] is frequently used to characterize the energetics of transitions in solid, fluid, and biomolecular systems. A routinely used technique to compute the PMF along a given reaction coordinate ξ is umbrella sampling. That technique aims to overcome limited sampling at energetically unfavorable configurations by restraining the simulation system with an additional (typically harmonic)

potential.[2] Accordingly, a set of N_w separate umbrella simulations are carried out, with an umbrella potential

$$w_i(\xi) = K_i/2(\xi - \xi_i^c)^2 \qquad (1)$$

which restrains the system at the position ξ_i^c ($i = 1, ..., N_w$) with a force constant K_i. From each of the N_w umbrella simulations (sometimes referred to as "umbrella windows"), an umbrella histogram $h_i(\xi)$ is recorded, representing the probability distribution $P_i^b(\xi)$ along the reaction coordinate biased by the umbrella potential $w_i(\xi)$. The probably most widely used technique to compute the PMF from histograms, that is, to unbias the distributions $P_i^b(\xi)$, is the weighted histogram analysis method (WHAM).[3]

On the basis of the histogram method of Ferrenberg and Swendsen,[4] the idea of WHAM is to estimate the statistical

---

* Author to whom correspondence should be addressed. Tel.: +46-(0)18-4715056. Fax: +46-(0)18-511755. E-mail: jochen@xray.bmc.uu.se.
† Uppsala University.
‡ Max-Planck-Institute for Biophysical Chemistry.

uncertainty of the unbiased probability distribution given the umbrella histograms, and subsequently to compute the PMF that corresponds to the smallest uncertainty. For a derivation of the equations, we refer to the original publication by Kumar et al.[3] An excellent (and less technical) review on umbrella simulations and the WHAM procedure has been presented by Roux.[5] The WHAM equations read[3]

$$P(\xi) = \frac{\sum_{i=1}^{N_w} g_i^{-1} h_i(\xi)}{\sum_{j=1}^{N_w} n_j g_j^{-1} \exp[-\beta(w_j(\xi) - f_j)]} \quad (2)$$

and

$$\exp(-\beta f_j) = \int d\xi \exp[-\beta w_j(\xi)] P(\xi) \quad (3)$$

Here, β denotes the inverse temperature $1/k_B T$, with the Boltzmann constant k_B and the temperature T, and n_j is the total number of data points in histogram h_j. The statistical inefficiency g_i is given by $g_i = 1 + 2\tau_i$, with the integrated autocorrelation time τ_i of umbrella window i (in units of the simulation frame time step.) Note that the g_i's cancel from the WHAM equations if (and only if) the autocorrelation times in all umbrella windows equal. In contrast, if the g_i's differ between different histograms, the factors g_i^{-1} assign lower weights to histograms with longer autocorrelations. $P(\xi)$ denotes the unbiased probability distribution that is related to the PMF via $\mathcal{W}(\xi) = -\beta^{-1} \ln[P(\xi)/P(\xi_0)]$. Here, ξ_0 is an arbitrary reference point where the PMF $\mathcal{W}(\xi_0)$ is defined to zero. The WHAM equations contain two unknown quantities, that is, the free energy constants f_j and the unbiased distribution $P(\xi)$, and must therefore be solved iteratively. Depending on the number of histograms and the height of the barriers in the PMF, the WHAM equations typically converge within tens of iterations and up to tens of thousands of iterations.

Alternative approaches to derive the PMF and the uncertainty from a set of umbrella simulations have been proposed,[6-8] as well as several extensions to the umbrella sampling technique.[9,10]

Despite the fact that WHAM has been widely used to derive PMFs from biomolecular simulations, a standard protocol to compute the statistical errors for the derived PMF has not yet evolved. Therefore, we here present a new WHAM implementation, termed g_wham, that allows one to compute robust error estimates using different bootstrap techniques. We apply the techniques on two test systems to demonstrate the potential and the limitations of the bootstrap methods. Besides the ability to estimate the statistical error, g_wham supports a number of features that are expected to be useful to the community. To compute PMFs along periodic reaction coordinates such as dihedral angles or coordinates in a simulation box with periodic boundary conditions, a periodic WHAM is implemented. Nonharmonic umbrella potentials can be provided as tabulated potentials. g_wham allows for the estimatation of autocorrelation times and the incorporation of these into WHAM. As shown in

the Results, this procedure may yield more realistic PMF estimates in the presence of long autocorrelations.

The software is freely distributed with the GROMACS simulation suite.[11] If the umbrella simulations were carried out using the GROMACS pull options, g_wham conveniently reads the GROMACS output files. In the case of more complex reaction coordinates, or if the simulations were not carried out using GROMACS, the user may provide g_wham input files in text format. A detailed description of g_wham, including all options, is provided in the Appendix and is available with the command line g_wham -h.

Methods

Error Estimates from Bootstrap Analysis. g_wham estimates the statistical uncertainty of the PMF using bootstrap analysis.[12] Bootstrapping is a resampling technique that can be applied to estimate the uncertainty of a quantity $A(a_1, ..., a_n)$ which is computed from a large set of n observations a_l ($l = 1, ..., n$). To calculate the uncertainty in A, one could redo the n observations multiple times, yielding several independent estimates for A and hence the uncertainty in A. That procedure would require many more observations and is therefore often not tractable.

The observations a_l are typically drawn from an unknown underlying probability distribution $P(a)$. The idea of bootstrapping is to estimate $P(a)$ using the n observations and subsequently generate new random sets of n hypothetical observations, based on the estimated distribution. Each of the sets of n hypothetical observations is used to calculate a hypothetical value for A. The uncertainty in A is then given by the standard deviation of the hypothetical values for A. For a detailed introduction into the bootstrap technique, we refer to the monograph by Chernick.[13]

Bootstrapping Trajectories Based on Umbrella Histograms. The WHAM procedure computes the PMF based on the N_w trajectories $\xi_i(t)$ along the reaction coordinate, each taken from one of the umbrella windows ($i = 1, ..., N_w$). All positions ξ_i during the N_w simulations may thus be considered as the large set of observations, which we referred to as a_l in the previous paragraph.[14] Alternatively, complete umbrella histograms may be considered as the individual observations (see next section).[15] Note that the probability distributions of ξ_i are already available as the umbrella histograms. Thus, we can generate new hypothetical observations, that is, a "bootstrapped" trajectory $\xi_{b,i}(t)$ for each umbrella histogram $h_i(\xi)$, such that $\xi_{b,i}(t)$ is distributed according to the respective histogram. Each bootstrapped trajectory $\xi_{b,i}(t)$ yields a new histogram $h_{b,i}(\xi)$. The new set of N_w histograms $h_{b,i}$ is subsequently applied in WHAM to compute a bootstrapped PMF $\mathcal{W}_b(\xi)$. The whole procedure is repeated N_b times (e.g., $N_b = 200$), yielding a large set of N_b bootstrapped PMFs $\mathcal{W}_{b,k}(\xi)$ ($k = 1, ..., N_b$). The uncertainty of the PMF is then given by the standard deviation as calculated by the N_b bootstrapped PMFs, that is via

$$\sigma_{PMF}(\xi) = \left[(N_b - 1)^{-1} \sum_{k=1}^{N_b} (\mathcal{W}'_{b,k}(\xi) - \langle \mathcal{W}'_b(\xi) \rangle)^2 \right]^{1/2} \quad (4)$$

Here, $\langle \mathcal{W}_b(\xi) \rangle = N_b^{-1} \sum_{i=k}^{N_b} \mathcal{W}_{b,k}(\xi)$ denotes the average of the bootstrapped PMFs at position ξ. One could also calculate the uncertainty via the standard deviation of the respective probabilities $\propto \exp(-\beta \mathcal{W}_{b,k}(\xi))$, which could subsequently be translated into the uncertainty of the PMF. We found that that this procedure yields similar error estimates compared to the definition in eq 4 applied here.

Any property generated from MD simulations has a natural time correlation. In order for the bootstrapping procedure to generate correct error estimates, that autocorrelation must be taken into account explicitly. Here, we chose the following procedure to generate autocorrelated bootstrapped trajectories $\xi_b(t)$ with a given integrated autocorrelation time (IACT) τ, and distributed according to a histogram $h(\xi)$. (Here, $h(\xi)$ may denote any of the given histograms, and the procedure is repeated for each histogram.) First, given a normally distributed random variable of zero mean and unit variance $R_t \sim \mathcal{N}(0,1)$, we generate a time series $x(t)$ via

$$x(0) = R_0 \tag{5}$$

$$x(t+1) = ax(t) + \sqrt{1-a^2}R_{t+1} \tag{6}$$

where $a = \exp(-1/\tau)$. Then, $x(t) \sim \mathcal{N}(0,1)$ and the IACT of $x(t)$ equals τ. The normally distributed $x(t)$ is translated into an evenly distributed series on $[0,1)$ using the error function via $x'(t) = (1 + \mathrm{erf}[x(t)\sqrt{2}])/2$. Eventually, we solve the equation

$$x'(t) = C_h(\xi_b(t)) \equiv \int_{-\infty}^{\xi_b(t)} h(\xi') \, \mathrm{d}\xi' \tag{7}$$

for $\xi_b(t)$, where $C_h(\xi_b(t))$ denotes the cumulative distribution function of the (normalized) histogram. Then, $\xi_b(t)$ will be distributed according to $h(\xi)$, with an approximate IACT of τ.

Bootstrapping Complete Histograms. The conformational sampling of macromolecules during MD simulations is frequently affected by long autocorrelations, with autocorrelation times ranging from pico- to microseconds or even longer. A complete sampling of all coordinates perpendicular to the reaction coordinate is therefore often intractable, in particular during a typically short umbrella simulation. In such situations, the individual umbrella histograms do not represent all accessible areas of phase space. Bootstrapped trajectories based on such nonconverged histograms, following the procedure in the previous paragraph, would also not represent all accessible areas of phase space. In addition, note that bootstrapping trajectories from given histograms require at least approximate knowledge of the IACT. Given only incomplete sampling, however, the IACT may be severely underestimated because slow transitions may not occur during the short umbrella simulations. Bootstrapping trajectories based on incomplete histograms in combination with underestimated IACTs would severely underestimate the uncertainty.

If the simulations are affected by such long autocorrelations, we suggest carrying out the simulation of each umbrella window multiple times from independent initial frames. Then, we consider complete histograms as individual observations and randomly select a new set of N_w histograms

from the given set of N_w histograms, allowing one to multiply select a specific histogram (sampling with replacement).[15] Hence, in contrast to the bootstrapping of trajectories based on umbrella histograms (see previous paragraph), we do *not* generate new trajectories and histograms. To ensure that the bootstrapped histograms span the whole reaction coordinate, that is, that no gaps between the bootstrapped histograms are generated, the histograms can be grouped along the reaction coordinate, and histograms can be bootstrapped within each group separately. We show that, given limited sampling, bootstrapping of complete histograms allows for a more accurate estimation of the uncertainty (see Results).

Bayesian Bootstrapping of Complete Histograms. As pointed out in the previous paragraph, introducing groups of histograms (and subsequent bootstrapping only within each group) avoids gaps along the reaction coordinate between bootstrapped histograms, but an appropriate choice for the number of histograms per group may be unclear. Therefore, we propose a method related to the so-called *Bayesian bootstrap* that avoids the introduction of groups of histograms by instead assigning random weights to all histograms within each bootstrap.

When applying the usual bootstrap on individual observations, n observations are selected with replacement from the given n observations a_i ($i = 1, ..., n$), where the probability of selecting any of the specific observations equals $1/n$. Hence, all observations a_i are selected with equal probability. Rubin proposed an alternative procedure, known as the Bayesian bootstrap, that instead assigns random weights ω_i to each observation.[16] Then, each observation a_i is selected with probability ω_i (instead of $1/n$), or alternatively, the weights ω_i are assigned to the observations when computing the observable $A(a_1, ..., a_n)$ from the observations. According to the Bayesian bootstrap, the weights ω_i are generated as follows: draw $n-1$ uniform random variables between 0 and 1, and let $u_{(1)}, u_{(2)}, ..., u_{(n-1)}$ denote their values in increasing order. In addition, let $u_{(0)} = 0$ and $u_{(n)} = 1$. The random weights are then given by the gaps between two consecutive random numbers, i.e., $\omega_i = u_{(i)} - u_{(i-1)}$, where $i = 1, ..., n$. For each bootstrap turn, new random weights are generated.

Note that the bootstrapping of complete histograms (compare previous section) is equivalent to the assignment of random weights to the histograms, if these random weights are an integer multiple of $1/N_w$. Here, we suggest the assignment of continuous random weights to the histograms, and selection of the weights according to the Bayesian bootstrap. That procedure resembles the bootstrapping of compete histograms in the sense that it considers only complete histograms as independent data points and thus is expected to yield realistic error estimates in the presence of long autocorrelations. However, because the continuous weights ω_i are (almost) never exactly zero, it excludes the possibility of generating gaps along the reaction coordinate in the bootstrapped histogram set. The WHAM procedure with weighted histograms was implemented by multiplying the inverse statistical inefficiencies g_i^{-1} in eq 2 by ω_i.

Autocorrelations. The normalized autocorrelation function of umbrella window i is given by

$$R_i(\Delta t) = \frac{\langle(\xi_i(t) - \langle\xi_i\rangle)(\xi_i(t + \Delta t) - \langle\xi_i\rangle)\rangle}{\sigma_{\xi,i}^2} \quad (8)$$

where $\xi_i(t)$ denotes the reaction coordinate during simulation i, $\sigma_{\xi,i}^2 = \langle(\xi_i(t) - \langle\xi_i\rangle)^2\rangle$ is the respective variance, and $\langle...\rangle$ represents the average over the simulation frames. Following the nomenclature in Kumar et al.,[3] the integrated autocorrelation time (IACT) of window i is defined by

$$\tau_{i,\text{int}} = \sum_{\Delta t=1}^{\infty} R_i(\Delta t) \quad (9)$$

The autocorrelation function derived from a short umbrella simulation is typically very noisy. Sophisticated methods to compute the IACT such as fitting of a single or double exponential to $R_i(\Delta t)$ or any kind of binning analysis turned out to be too unstable for the present purpose. Note that the IACT should be computed automatically for hundreds (or thousands) of possibly poorly converged $R_i(\Delta t)$'s. Therefore, we chose to compute $\tau_{i,\text{int}}$ directly via eq 9 but carried out the summation only until $R_i(\Delta t)$ dropped under a predefined threshold of 0.05.

Simulation Details. Test simulations were carried out using the GROMACS simulation suite.[11] As a test system, we have computed the PMF along the distance between two methanol molecules in a vacuum. These simulations were set up by placing one methanol molecule in the origin and placing the second molecule at the distance ξ_i^c of the corresponding umbrella window. The molecules were randomly rotationally oriented. In addition, the initial distance between the two molecules was varied randomly by $\pm\sigma_u$, where $\sigma_u = \sqrt{(k_BT/K)}$ denotes the width of the umbrella histogram (assuming a flat underlying PMF). Here, $K = 800$ kJ/mol/nm^2 is the umbrella force constant, k_B is the Boltzmann factor, and T is the temperature. The sampling was carried out using a stochastic dynamics integrator ($\tau = 0.07$ ps, $T = 300$ K), with an independent random seed for each simulation. Lennard-Jones and electrostatic interactions were computed in direct space without a cutoff. Bonds were constrained using LINCS,[17] allowing a time step of 2 fs. The umbrella positions were recorded every 10th step during simulations with a total simulation time of 50 ps and were recorded at every step during simulations with a simulation time of 4 ps. Methanol parameters were taken from the GROMOS96 force field.[18] The methods applied to compute the PMF for the Rhesus channel Rh50 and for the lipid membrane have been published elsewhere.[19]

Results and Discussion

Error Analysis for a Model System and a Lipid Membrane PMF. As a test system, we compute the PMF of the center-of-mass distance between two methanol molecules in a vacuum. Such a simple system allows us to carry out the complete set of umbrella simulations many times and hence to accurately compute the "true" statistical uncertainty of the PMF. Subsequently, we test whether the bootstrapping
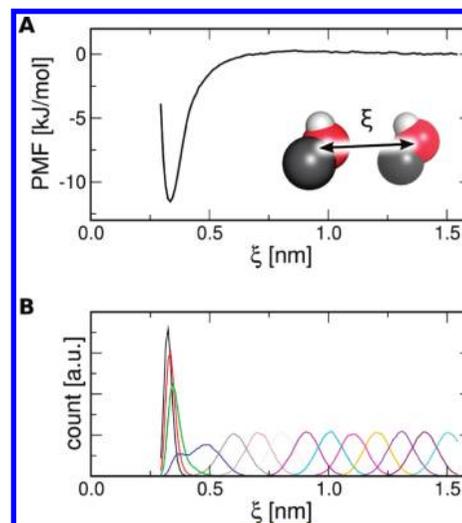

Figure 1. (A) Converged PMF (black curve) of the center of mass distance between two methanol molecules in vacuum. (B) Converged umbrella histograms, each derived from a 3-ns simulation.

procedures are able to estimate the "true" uncertainty using only the data from one set of umbrella simulations.

As a reference for the following discussion, Figure 1A presents the converged PMF and Figure 1B, the respective umbrella histograms. Here, each of the 14 umbrella windows was simulated for 3 ns, yielding well-converged statistics as visible from the Gaussian histograms at a great distance ξ. The PMF at $\xi = 1.5$ nm was chosen as a reference point and defined to zero. To arrive at a flat PMF at a great distance, the PMF was corrected by $k_BT \ln(4\pi\xi^2)$, which removes the entropic decrease in the PMF because of the increase in the number of configurations on a sphere of radius ξ.[20]

To assess whether the bootstrapping procedure provides a reliable error estimate, we have repeatedly computed the same PMF using limited statistics, with each umbrella window simulated for 50 ps. The nonconverged histograms of one set of these umbrella simulations is shown in Figure 2A. The complete set of umbrella simulations was carried out 50 times with different initial random seeds for the stochastic forces and different initial orientations and velocities of the methanol molecules, yielding 50 independent estimates for the PMF (Figure 2B). The uncertainty (67% confidence interval) for a single set of umbrella simulations as derived from these 50 PMFs is plotted as a green curve in Figure 2D. Note that the error at $z = 1.5$ nm equals zero since all PMFs were defined to zero at that point. Next, an autocorrelated bootstrapped trajectory was generated for each of the histograms plotted in Figure 2A using eqs 5, 6, and 7, allowing one to compute a new "hypothetical" estimate for the PMF based on the umbrella histograms. That bootstrapping procedure was repeated 200 times, yielding 200 bootstrapped PMFs (Figure 2C, colored curves). As expected, the bootstrapped PMFs substantially differ, in line with the 50 PMFs calculated from independent simulations (Figure 2B). The standard deviation computed from the bootstrapped PMFs (Figure 2D, black) is in good agreement with the uncertainty calculated from the 50 independent
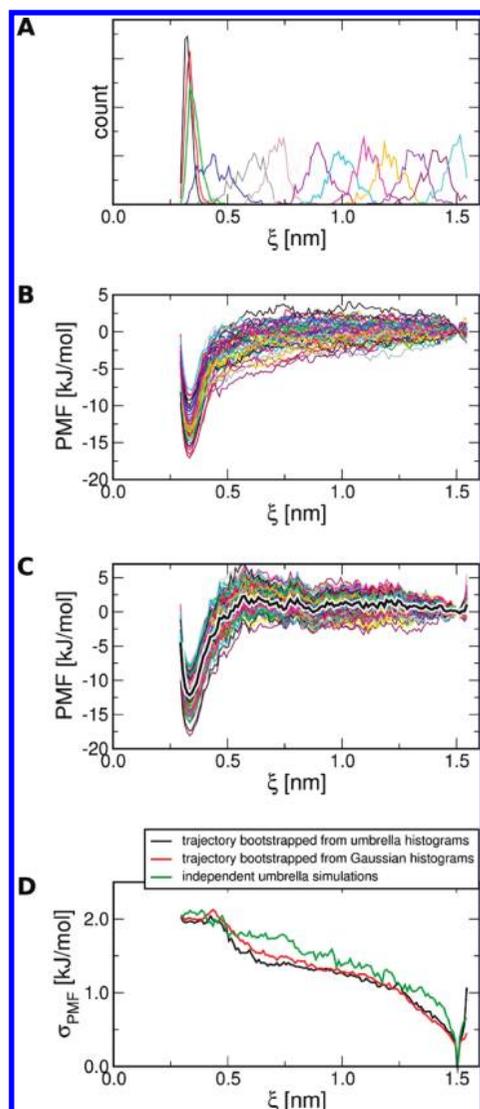
g_wham−A Free WHAM Implementation

J. Chem. Theory Comput., Vol. 6, No. 12, 2010 **3717**

Figure 2. (A) Nonconverged histograms, each derived from 50 ps simulations. (B) 50 PMFs derived from 50 fully independent sets of umbrella simulations. (C) PMF (black curve) derived from the set of nonconverged histograms (A). Autocorrelated trajectories were bootstrapped from the histograms shown in A 200 times, yielding 200 bootstrapped PMFs (colored curves in C). (D) Statistical uncertainty calculated from the 50 independent simulations (green) shown in B and from the 200 bootstrapped PMFs (black) shown in C. Alternatively, the uncertainty was estimated from trajectories that were bootstrapped from Gaussian distributions of the average and σ taken from the umbrella histograms (red).

Figure 3. Estimating uncertainties in the presence long autocorrelations. (A) The PMF along the methanol−methanol distance (not shown) was computed from 140 umbrella histograms, each derived from a 4 ps simulation. As a reference, the uncertainty σ_{PMF} was computed from 100 independent sets of umbrella simulations (green curve). Generating bootstrapped trajectories for each umbrella histogram leads to an underestimated uncertainty (black curve). Estimating the uncertainty by bootstrapping complete histograms (red curve) or using the Bayesian bootstrap on complete histograms (blue curve) yields more accurate error estimates. (B) PMF for ammonia permeation across a lipid membrane containing 40 mol % cholesterol. (C) Statistical uncertainty of the ammonia PMF computed by bootstrapping trajectories for each umbrella histogram (black curve) and by (Bayesian) bootstrapping of complete histograms (red and blue curves).

simulations (Figure 2D, green), demonstrating that the bootstrapping procedure provides a reliable error estimate without the requirement to carry out new independent simulations. Alternatively, the uncertainty was estimated from trajectories that were bootstrapped from Gaussian distributions with the average and width taken from the respective umbrella histogram (Figure 2D, red), yielding almost identical and hence equally accurate error estimates.

   Biomolecular simulations naturally contain long autocorrelations. The histograms based on short umbrella simulations may therefore not represent all parts of phase space. In addition, the IACTs may be severely underestimated since slow transitions do not occur during the short simulations.
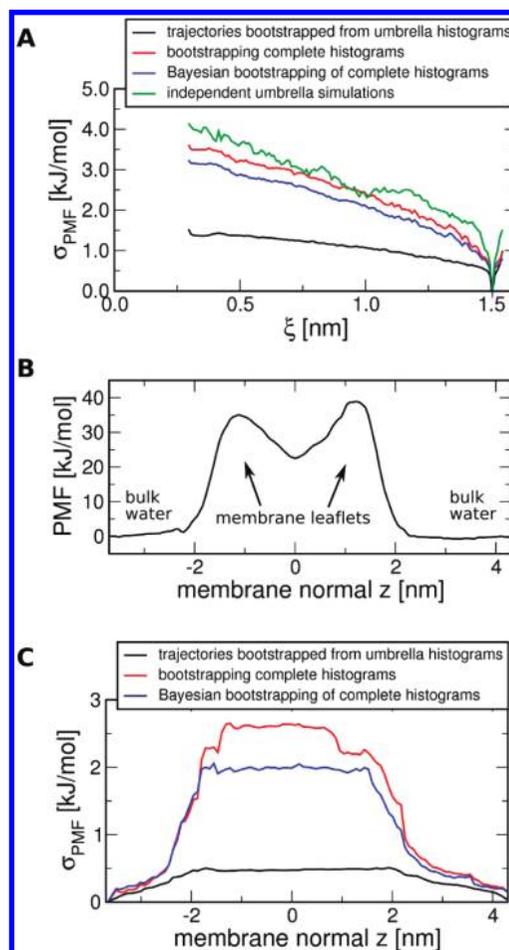
Consequently, bootstrapping trajectories based on these histograms (in combination with underestimated IACTs) will underestimate the uncertainty. This fact is demonstrated in Figure 3A. To emulate umbrella sampling of a biomolecular system with long autocorrelations, we computed the PMF of the methanol distance based on 4 ps simulations (using the first 0.5 ps for equilibration), resulting in highly nonconverged histograms. Ten independent umbrella simulations were carried out for each of the 14 umbrella window positions, yielding 140 histograms. The whole set of umbrella simulations was carried out 100 times, allowing one to compute the true uncertainty (as one standard deviation) in the PMF (Figure 3A, green curve). Figure 3A compares the

3718 *J. Chem. Theory Comput., Vol. 6, No. 12, 2010*

Hub et al.

true uncertainty to the estimated uncertainty derived from three different bootstrapping methods. Because the estimated uncertainties vary slightly between the different sets of independent umbrella simulations, Figure 3A plots estimated uncertainties averaged from 15 (of the 100) sets of umbrella simulations. The uncertainty computed by bootstrapping trajectories is shown as a black curve, demonstrating that this procedure greatly underestimates the uncertainty in that case. The red curve in Figure 3A presents the uncertainty estimated by bootstrapping complete histograms. Here, the histograms were grouped into 14 sets of 10 histograms, with each group containing the 10 histograms at the same umbrella position. Consequently, 10 histograms were bootstrapped from each of the 14 sets, and the PMF was computed from the 140 bootstrapped histograms using WHAM. The whole procedure was repeated 200 times, providing 200 hypothetical estimates for the PMF (not shown) and allowing one to compute the uncertainty using eq 4. As visible from Figure 3A, bootstrapping complete histograms yields a more accurate estimate of the uncertainty, despite the poor sampling within each umbrella window. The blue curve in Figure 3A presents the uncertainty estimated using Bayesian bootstrapping of complete histograms, that is, by assigning random weights to the individual histograms (see Methods). The Bayesian bootstrap also yields a reasonable error estimate because the method considers only complete histograms as independent data points, similar to the bootstrapping of complete histograms.

For a second comparison between the different bootstrapping methods, Figure 3B presents the PMF for ammonia permeation across a biological membrane composed of the lipids POPE and POPC plus 40 mol % cholesterol. The flat regions at small and large z correspond to the ammonia molecule in the two bulk water regions above and below the membrane, whereas the two maxima in the PMF correspond to the hydrophobic regions of the two membrane leaflets. The PMF has been computed from 656 histograms (not shown), each taken from 1 ns of simulation, where the first 50 ps were removed for equilibration. The initial frames for the umbrella simulations at a specific z coordinate were generated by inserting ammonia at various randomly chosen positions in the membrane plane, justifying the assumption that the histograms are independent. Figure 3C shows the estimated uncertainty computed via (i) bootstrapped trajectories (black), (ii) bootstrapping of complete histograms with 12 histograms within each group (red), and (iii) Bayesian bootstrapping of complete histograms (blue). Method i yields a very small uncertainty of only 0.5 kJ/mol, whereas methods ii and iii yield an uncertainty of ~2 kJ/mol at the main barriers in the PMF. Because considerable computational effort is required to compute the PMF in Figure 3B, we cannot compute the uncertainty from independent sets of umbrella simulations for this example. However, Figure 3C suggests that the individual histograms do not represent all accessible areas of phase space, leading to an underestimated uncertainty as computed from method i. Presumably, slow transitions on a multi-nanosecond time scale may affect the sampling in this case, whereas the autocorrelation analysis based on the shorter simulations yields spuriously short

IACTs. In contrast to method i, methods ii and iii do not depend on the accurate computation of the IACTs but only require the histograms to be independent. Therefore, methods ii and iii are expected to yield a reliable error estimate in this case.

To estimate uncertainties in the presence of long (possibly unknown) autocorrelations, we therefore suggest carrying out many short umbrella simulations instead of a few long umbrella simulations, such that each position along the reaction coordinate is covered by at least several independent histograms. Given sufficiently many independent histograms, the error can be estimated using bootstrapping of complete histograms or using the Bayesian bootstrap of complete histograms.

Effect of Autocorrelations. As visible from the WHAM equations, eqs 2 and 3, the IACTs cancel if (and only if) the IACTs are equal in all umbrella windows. In nonhomogeneous systems, however, that assumption may not hold. An example would be umbrella simulations for solute permeation across a lipid membrane or across a protein channel surrounded by bulk water. Here, the IACTs of windows in the bulk are typically lower than the IACTs of windows inside the lipid membrane or inside the protein channel. We found that neglecting the IACTs may lead to artifacts in particular when computing the PMF along a periodic reaction coordinate. As an example, Figure 4A presents a nonconverged PMF for ammonia permeation across the Rhesus protein channel Rh50 from *N. europaea* (Figure 4C). The PMF was derived from 365 400-ps histograms, taken from 500 ps simulations, using the first 100 ps for equilibration. The simulations were carried out with periodic boundary conditions, implying that a PMF for solute permeation should yield the same free energy in the two bulk-water regions below and above the channel. The black curve in Figure 4A was computed by a nonperiodic WHAM. The PMF is not converged, as apparent from the substantial offset of ~15 kJ/mol between the two bulk-water regions. To account for the periodicity of the system, a periodic WHAM assuming equal IACTs of all umbrella windows could be carried out (red curve). However, with equal IACTs, the WHAM procedure assigns equal weights to all histograms and, hence, equally distributes the offset of 15 kJ/mol along the reaction coordinate to enforce a periodic PMF. As a consequence, an unphysical slope is induced in the bulk-water regions of the PMF ($|z| > 2$ nm). A more realistic procedure is therefore to compute the IACTs for each umbrella window and to apply them within WHAM. The IACT derived by direct integration of the autocorrelation function for the displacement for each umbrella window is plotted in Figure 4B as black dots. Because the IACTs cannot be accurately computed from the limited sampling in the umbrella windows, we suggest smoothing the IACT along the reaction coordinate yielding a semiquantitative autocorrelation measure (Figure 4B, red curve). Whereas the IACTs are small in bulk water, substantial autocorrelations limit the sampling within the channel, suggesting that the 15 kJ/mol is a consequence of slow sampling within the channel. The PMF computed by a periodic WHAM that takes IACTs into account is shown in Figure 4A as a blue curve. As expected, the PMF is flat in the bulk-water regions (in agreement with the nonperiodic WHAM result, black curve), whereas corrections
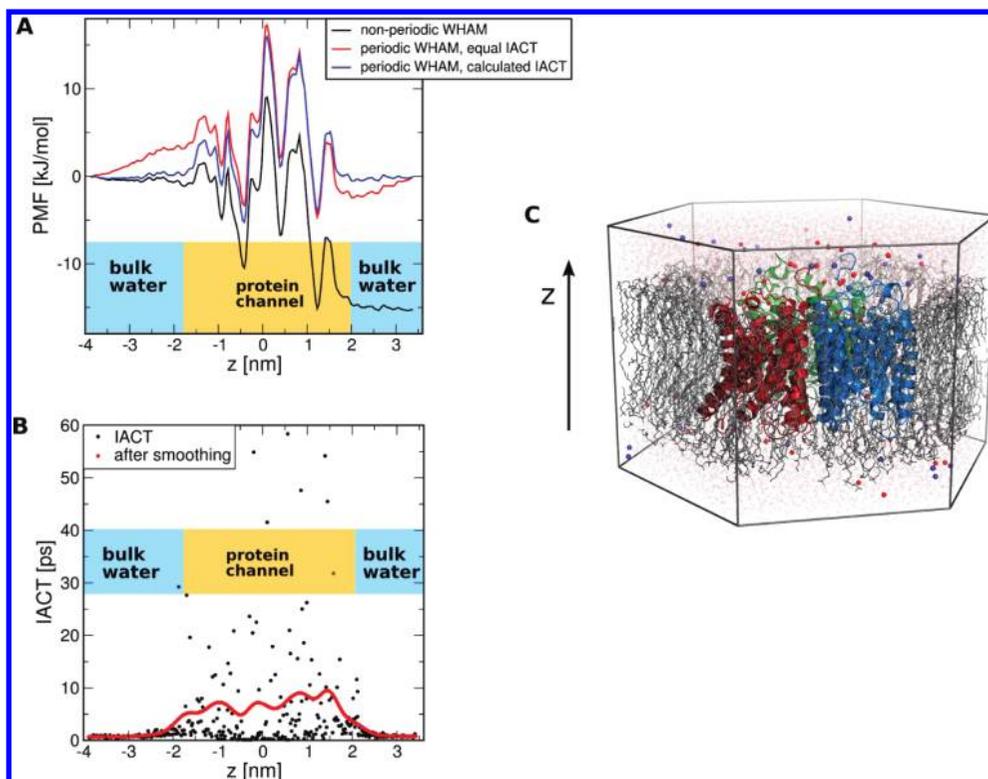
g_wham—A Free WHAM Implementation

J. Chem. Theory Comput., Vol. 6, No. 12, 2010 **3719**

Figure 4. Effect of autocorrelations in a periodic WHAM. (A) Nonconverged PMF of ammonia permeation across the Rhesus protein channel Rh50 (black). The limited sampling accounts for a substantial offset of ~15 kJ/mol between the two end points of the PMF corresponding to the two bulk water regions. A periodic WHAM assuming equal integrated autocorrelation times (IACTs) accounts for the periodicity of the system (red curve) but induces approximately a linear slope in the complete PMF, including the well-sampled bulk water regions. Blue curve: PMF derived from periodic WHAM incorporating the calculated IACTs. The PMFs in the bulk-water regions are almost flat, in accordance with the bulk-water regions in the nonperiodic PMF (black). (B) IACTs calculated by direct integration of the autocorrelation functions (black dots), and by subsequent smoothing with a Gaussian filter (red curve). (C) Simulation box of an Rh50 trimer embedded in a lipid membrane and solvated in water and 150 mM electrolyte.

were introduced in the less sampled channel region to yield a periodic PMF.

Converged PMFs for ammonia permeation across the Rh50 channel as well as the biological implications have been published elsewhere.[19]

Conclusions

We have presented a new WHAM implementation, termed g_wham, that is freely distributed with the GROMACS simulation suite. The g_wham software is easy to use, flexible, and efficiently implemented. Statistical uncertainties are quantified using different bootstrap analysis methods: (i) bootstrapping of hypothetical trajectories based on the umbrella histograms together with the respective autocorrelation time, (ii) by bootstrapping complete histograms, or (iii) by using the Bayesian bootstrap of complete histograms, that is, by assigning random weights to the histograms. We have shown that method i provides an accurate error estimate if (and only if) the histograms are sufficiently converged. If the histograms are affected by long autocorrelations, as frequently occurrs in simulations of large biomolecules, methods ii and iii provide a more accurate error estimate. In nonhomogeneous systems such as a protein channel or a lipid membrane surrounded by bulk water, the autocorrelation times may substantially vary along the reaction coordinate

and thus not cancel from the WHAM equations. Consistent application of the autocorrelations has here been shown to yield a more accurate estimate for the PMF in such systems, in particular when computing a periodic PMF.

Appendix

g_wham Input Modes. A help file, including all command line options, is provided by the g_wham tool via the command g_wham -h. g_wham supports three input modes. In modes 1 and 2, g_wham reads specific GROMACS files, These modes are thus convenient for GROMACS users. In mode 3, g_wham reads only text files and is therefore suitable for non-GROMACS users as well.

1. With option -it, the user provides a file which contains the *file names* of the umbrella simulation run-input files (GROMACS tpr-files). In addition, with option -ix, the user provides a file which contains the file names of the pull position output files (pullx.xvg etc.) written by the GROMACS mdrun program.

2. This mode is the same as mode 1, except that the user provides with option -if a file which contains the file names of the pull force output files (pullf.xvg etc.) written by the GROMACS mdrun program.

3. With option -ip, the user provides a file which contains the file names of the pull output files written by GROMACS 3 (pdo files). pdo files are text files and can be generated by non-GROMACS users. Each pdo file contains a header with the umbrella positions and force constants, and the body contains the simulation time versus the displacement of the system with respect to the umbrella center. The pdo file format (with a typical header) is explained with the g_wham help file provided with g_wham -h.

WHAM Options. Default values for the following options are listed in square brackets:

-min, -max: boundaries of the profile [0,0]

-auto: determine boundaries automatically [yes]

-bins: number of bins used [200]

-temp: temperature in Kelvin [298.15]

-tol: tolerance. The WHAM iterations stop when the probabilities change less than the tolerance. [10^{-6}]

-b, -e, -dt: specify simulation times in picoseconds (begin, end, time step) that are used in WHAM [50, infinity, 0]

-cycl: periodic (or cyclic) WHAM [no]

-tab: file name with tabulated potential in the case of nonharmonic umbrella potentials

Output Control

-o: file name of PMF output file

-hist: file name of histogram output file

-histonly: write histograms and exit [no]

-boundsonly: determine boundaries automatically and exit [no]

-log: write negative logarithm of the probabilities; that is, enable output in energy units; otherwise, write probabilities [yes]

-unit: define energy unit (kJ/mol, kcal/mol, k_BT) [kJ/mol]

-zprof0: set profile to zero at this position [0]

-sym: symmetrize profile around $\xi = 0$ (useful for membranes, for instance) [no]

-v: verbose mode [no]

Autocorrelation Handling

-ac: calculate integrated autocorrelation times (IACTs) using eqs 8 and 9 and use in WHAM [no]

-acsig: smooth IACTs along reaction coordinate using a Gaussian filter of width defined here [0]

-ac-trestart: when computing the autocorrelation functions for $\xi_i(t)$, restart the calculation after the time delay defined here [1 ps]

-oiact: (smoothed) IACT output file name

-iiact: IACT input file name. If the user prefers to calculate the IACTs *not* using g_wham, the IACTs can be provided to g_wham using this option.

Bootstrapping Control

-bsprof: output file name of all bootstrapped profiles

-bsres: output file name with average and standard deviation of bootstrapped profiles (that is, the uncertainty of the PMF)

-nBootstrap: number of bootstraps carried out to estimate the uncertainty (use, e.g., 100) [0]

-bs-method: bootstrap method applied ('b-hist', 'hist', 'traj', or 'traj-gauss'); Bayesian bootstrapping of complete histograms, bootstrap complete histograms, bootstrap new trajectories from the umbrella histograms, or bootstrap new trajectories from Gaussian distributions with average and width taken from the respective histogram [b-hist]

-bs-tau: specify integrated autocorrelation time used for all histograms with bootstrap methods 'traj' or 'traj-gauss'; if not provided (default), use calculated IACTs (options -ac and -acsig)

-histbs-block: number of histograms in one group with bootstrap method 'hist'; histograms will be bootstrapped only within each group separately; that procedure avoids gaps without any histogram data along the reaction coordinate.

-bs-seed: random seed for bootstrapping (−1 generates a seed) [−1]

-vbs: verbose bootstrapping (output cumulative distribution functions for each histogram and a histogram file for each bootstrapped PMF) [no].

References

(1) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(2) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578–581.

(3) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(4) Ferrenber g, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.

(5) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.

(6) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104.

(7) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *124*, 234106.

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

(9) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865–880.

(10) Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40–57.

(11) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(12) Efron, B. *Ann. Stat.* **1979**, *7*, 1–26.

(13) Chernick, M. R. *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed.; Wiley-Interscience: New York, 2007.

(14) Grossfield, A. An implementation of WHAM: the weighted histogram analysis method. http://membrane.urmc.rochester.edu/Software/WHAM/WHAM.html (accessed October 6, 2010).

(15) Hub, J. S.; de Groot, B. L. *Biophys. J.* **2006**, *91* (3), 842–848.

(16) Rubin, D. B. *Ann. Stat.* **1981**, *9*, 130–134.

(17) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.

(18) Van Gunsteren, W. F.; Berendsen, H. J. C. *Gromos Manual*; BIOMOS, Biomolecular Software, Laboratory of Physical Chemistry, University of Groningen: Groningen, The Netherlands, 1987.

(19) Hub, J. S.; Winkler, F. K.; Merrick, M.; de Groot, B. L. *J. Am. Chem. Soc.* **2010**, *132*, 13251–13263.

(20) Neumann, R. *Am. J. Phys.* **1980**, *48*, 354–357.