

Absolute entropies from molecular dynamics simulation trajectories

Heiko Schäfer, Alan E. Mark, and Wilfred F. van Gunsteren

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH Zentrum,
CH-8092 Zürich, Switzerland

(Received 20 March 2000; accepted 21 July 2000)

A heuristic formula for calculating absolute entropies from the covariance matrix of atom-positional fluctuations was extensively tested. Because of its heuristic nature, the results obtained are compared to analytical expressions for an ensemble of harmonic oscillators, for the ideal gas, and to numerical results obtained from the equation of state for the Lennard-Jones fluid as a means of validation of the approximate formula for the entropy. The formula yields rather accurate results. The removal of translational and rotational rigid body motion and the effect of the various fitting procedures involved are discussed for the more realistic system of a β -heptapeptide in solution.

© 2000 American Institute of Physics. [S0021-9606(00)51539-6]

I. INTRODUCTION

Molecular dynamics (MD) simulations are a well-established tool to investigate the stability and behavior of systems ranging from abstract models to complex molecular assemblies of biological interest.¹⁻⁵ The energetics of these systems is easily accessible using MD and even relative free energies are now routinely calculated, albeit at a high computational cost. However, the calculation of entropies remains complicated because the entropy depends on the whole phase space of the system of interest.⁶ This is true even for entropy differences.

Most methods^{7,8} for entropy calculations based on MD simulations have been restricted to the configurational entropy of a molecule as the sampling of the translation and rotation of the molecule was not complete. These methods used a transformation to internal coordinates to separate the slow overall motions from the comparably fast internal motions. Complications due to the handling of constraints were usually neglected. To calculate entropies the probability distribution of each degree of freedom (DOF) must in principle be known. However, it is notoriously difficult to sample probability distributions, especially their tails. A typical approximate solution to this problem is the assumption of a particular functional form for the probability distribution. Karplus and Kushick⁷ used a Gaussian form for each internal degree of freedom (bonds, bond angles, dihedrals) and the correlation between them. Di Nola *et al.*⁸ tried a combination of direct sampling for each degree of freedom and a Gaussian approximation for the correlation. They found that the inclusion of correlations was important. The Gaussian approximation to the probability distribution of each degree of freedom seemed to be appropriate.

In 1993, Schlitter⁹ proposed yet another approximate method that is rationalized using the Gaussian approximation to the probability distribution but elegantly circumvents the need to express the entropy in internal coordinates. Schlitter's formula is of heuristic nature and constitutes an upper bound to the entropy. The approximate formula was rationalized by showing that the correct quantum and classical limits

are obtained for the temperature approaching zero or infinity.⁹ The numerical tests presented in Ref. 9 were, unfortunately, of limited value. Here, Schlitter's heuristic formula is extensively tested with the aim of its validation. Second, as the method allows us to use Cartesian coordinates and does not restrict us to the configurational entropy, it is used to investigate entropic effects in the reversible folding of a small peptide.⁵ We have implemented Schlitter's formula to work with the GROMOS96 simulation package.^{10,11}

II. METHOD

Schlitter⁹ introduced a very elegant formula to calculate absolute entropies from MD trajectories using the covariance matrix of atom-positional fluctuations. In this section we summarize the heuristic derivation of Schlitter's formula in order to clarify the approximations on which it is based.

The formula is based on a quantum-mechanical treatment of a one-dimensional degree of freedom x with states $|n\rangle, n=0,1,2,\dots$. The energy of each state is ϵ_n , the mass of x is m , and its mean $\langle x \rangle$ is assumed to be zero. The canonical partition function at temperature T is

$$Z = \sum_n \exp(-\beta \epsilon_n), \quad (1)$$

with $\beta = 1/k_B T$, the Boltzmann constant k_B and temperature T . The complete entropy S of the system can be expressed using the probability $p_n = \exp(-\beta \epsilon_n)/Z$ of finding the system in state n as

$$S = -k_B \sum_n p_n \ln p_n. \quad (2)$$

For a given expectation value (ensemble average) of the variance $\langle x^2 \rangle = \sum_n p_n \langle n | x^2 | n \rangle$, the entropy is maximized by varying the probabilities p_n . The variance $\langle x^2 \rangle$ and the normalization of the probabilities $\sum_n p_n = 1$ are taken into account as constraints using Lagrange multipliers.

This results in the demand that the energy of each state must be proportional to the state's variance: $\epsilon_n \propto \langle n | x^2 | n \rangle$. A system for which this condition is fulfilled is the simple har-