JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

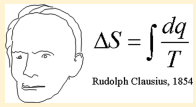
Calculation of Molecular Entropies Using Temperature Integration

Kim Sharp*

E. R. Johnson Research Foundation, Dept. of Biochemistry and Biophysics, The Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059, United States

Supporting Information

ABSTRACT: There are two fundamental definitions of entropy, Clausius's thermodynamic definition and the Boltzmann–Gibbs statistical mechanical definition. The Clausius definition, applied here to the calculation of molecular entropies, requires an integration over temperature. Analytical and numerical error analysis shows that this quadrature can be done accurately using a small number of large temperature steps involving the calculation of one average system property, the mean internal energy. This makes the method computationally practical for systems with many degrees of freedom, such as biological molecules. The Clausius definition provides a simple and physically insightful way to decompose a total entropy change into components and is useful for benchmarking statistical mechanical based methods for entropy calculation. A temperature quenching protocol is described whereby the Clausius method can be used with existing force fields to evaluate entropy changes in anharmonic and diffusive systems.



INTRODUCTION

The change in entropy (S) plays a crucial role in many chemical and biological reactions through its contribution to the free energy change, the other contribution being the enthalpy/energy. The primacy of free energy relative to entropy and enthalpy is that it determines the equilibrium constant by $K = \exp(-\Delta G/kT)$, where k is Boltzmann's constant, T is temperature, and ΔG is the Gibbs free energy. The working assumption of many experimental and theoretical biophysicists is that knowledge of the separate entropy and enthalpy contributions can provide additional "mechanistic" insight into the relationship between the structure and dynamics of the molecules involved and their function. This is particularly helpful for biological systems due to the complexity of the molecules involved and the large number of degrees of freedom. In allostery, for example, two conceptually different mechanisms have been proposed. In the Monod–Wyman–Changeux type model, coupling between distal sites occurs through a shift between different (average) structures.[2] In contrast, the Cooper–Dryden model allows for an entirely entropic mode of coupling between effector binding and protein response with no change in average structure.[3] On the experimental side, NMR relaxation studies show that entropy changes play a controlling role in recognition of many different peptides by the protein calmodulin.[4,5] In drug development, a common design strategy is to rigidify the ligand in order to minimize the entropy loss upon binding to its target.[6] There is evidence that historical development of drugs within a class follows a progression from entropy dominated to enthalpy dominated interactions.[7] Consequently, there is very active development of practical methods for calculating entropy changes directly from theory and simulation. This approach is distinct from, but complementary to, methods such as thermodynamic integration and free energy perturbation. These methods are designed to compute free energy changes. The entropy change can then be obtained indirectly by numerical differentiation with respect to T.

Direct calculation of entropy changes is based on the Boltzmann–Gibbs–Planck (BGP) expression for entropy[8]

$$S = -k \int p(\mathbf{q}) \ln p(\mathbf{q})\, d\mathbf{q} \quad (1)$$

where $\mathbf{q}$ is the vector of N coordinates describing the position of all the atoms. In coarse grained representations, $\mathbf{q}$ may represent coordinates of aggregates of atoms. $p(\mathbf{q})$ is the N-dimensional probability distribution function (pdf), and integration is taken over the entire positional phase volume. Accurate evaluation of eq 1 is challenging. For any system of biological interest, the dimensionality is high: O(100) or more. This presents formidable sampling difficulties. One strategy is to benchmark a computational method on a low-dimension system (~O(10)) with either an analytically known entropy or one that can be calculated by brute force to high numerical accuracy. This still requires an extrapolation of uncertain accuracy to the much larger biological system of interest. Brute force evaluation is therefore problematic. Another strategy is to compare different approximations to eq 1, taking advantage of the plethora of alternative methods that have been developed, which include harmonic approximations, quasi-harmonic approximations, quasi-harmonic approximations with anharmonic and higher order correlation corrections, methods based on counting of conformational states, and important states (i.e., low energy state) sampling.[9−25]

An expression for "information" entropy identical to eq 1 except for the constant k arises from information theory.[26,27] This measure is widely used in applications such as pattern recognition and signal processing. These fields therefore face the same difficulties in evaluating a high dimension pdf.

Received: October 17, 2012
Published: December 18, 2012

Consequently, methods developed in the field of information theory are directly applicable to the calculation of thermodynamic entropies. A good example is the use of the mutual information expansion (MIE) of the entropy.[21,28]

$$S = S_1 + ... + S_N - I_{1,2} - I_{1,3}...I_{i,j>i} + I_{1,2,3} + ...I_{i,j>i,k>j} + ...I_{\{1...N\}} \quad (2)$$

where $S_i/k = -\int \delta q_i\ p(q_i)\ \ln\ p(q_i)$ is the entropy from coordinate q_i alone (neglecting all correlations) obtained by integrating over the one-dimensional (1-D) marginal pdf of q_i. $I_{\{i...n\}}$ is the mutual information shared between the subscripted coordinates due to correlation between their distributions. For example, the second order terms are given by $I_{ij} = -k\int \delta q_i\ \delta q_j\ p(q_iq_j)\ \ln\ p(q_iq_j)/(\ln\ p(q_i)\ln\ p(q_i))$ which involve only 2-D and 1-D pdf's. Generally, there will be $N!/(n!(N-n)!)$ mutual information terms of order n involving n-dimensional and lower pdf's, culminating in the final Nth order term $I_{\{1...N\}}$ involving the N-D pdf. In applications to biomolecules,[21,24] inclusion of second and third order terms improves on the naïve estimate of no correlation ($S = \sum S_i$) using only 1-D pdf's. In principle, inclusion of higher and higher order terms will further improve the entropy estimate, but the current practical limit for sampling and evaluation for medium to large biomolecules appears to be 3-D pdf's.[24] For example, if one considered just the torsional degrees of freedom of a small protein with say 120 torsion angles binned in 5° increments, a third order MIE already requires more than a quarter-million 3-D pdf's each with more than 1.7 million data elements. The MIE, while insightful due to the information theory interpretation of its terms, is not a unique expansion of the entropy, nor are there *a priori* reasons why MIE should be optimal computationally. Whether the magnitude of MIE terms generally decreases with increasing order, the extent to which various terms tend to cancel, and what order is required for numerical convergence are open questions. Indeed, Tidor et al. have recently applied the maximum information spanning tree (MIST) approach to this problem specifically to identify and evaluate the larger magnitude terms of the entropy expansion more efficiently.[29] In addition to the issues of convergence and computational demand, there is that of bias from truncation of the expansion: A survey of different information expansion estimators suggests that there are no unbiased methods.[30]

The very diversity of methods currently being used to approximate and evaluate eq 1 indicates that computation of the entropy in biological systems remains challenging. However, the biophysicist has one advantage over the information theorist: Thermodynamic entropy has another definition,[31] which in fact predates the statistical mechanics definition by some 20 years:

$$\Delta S = \int \frac{dq}{T} \quad (3)$$

The integration is taken over some reversible path from initial to final states, and dq is the heat entering the system (if positive) or leaving (if negative). It is remarkable that eqs 1 and 3, though so different, give exactly the same entropy change for an equivalent process. There appears to be no information theory counterpart to the thermodynamic quantities of heat and temperature, and so no counterpart of eq 3.

The goal of this paper is to explore the advantage the Clausius expression gives when calculating a thermodynamic entropy. First, the feasibility of accurately evaluating the heating integral is demonstrated. Then, various applications of the method are described: providing a physical interpretation of entropy changes, benchmarking other entropy calculation methods, and calculating entropy changes in diffusive systems and in multiscale systems.

THEORY AND METHODS

Preliminary Remarks. The heat referred to in eq 3 is defined as the difference between the change in internal energy of the system and the work done by the system, namely, $dq = dE - w$. I consider here typical biological conditions where $P = 1$ atm, so PdV work is negligible. Then, to a good approximation, the heat entering the system is equal to the change in total internal energy (dE), which is the sum of the kinetic and potential energies. Under these conditions, $dG \approx dA$, the change in Helmholtz free energy. Extending the results of the present analysis to cases where significant PdV or other external work is done is straightforward, requiring only the more general equation for heat transfer, above.

Changes in equilibrium conditions modeled in simulations of the type discussed in the present work are *de facto* reversible in that regeneration of the same macroscopic state (specified by T, V, N) involves no irreversible change in any external environment. Thus, eq 3 is applicable to the calculation of entropy changes.

Note that eq 1 describes only the positional contribution to the entropy in the classical approximation. Due to the additivity of the positional and momentum contributions to the Hamiltonian in a classical treatment, a separate integration over the particle momenta $\mathbf{p}$ provides the kinetic contribution to the entropy. The counterpart for eq 3 is that the heat entering the system alters both kinetic energy and potential energy, with corresponding changes in the kinetic entropy and positional entropy. Whether using eq 1 or eq 3, the kinetic contribution to entropy (and other thermodynamic quantities) is separable and trivially known from the temperature, atomic masses, and equipartition. It carries no useful information for the biophysicist and is easily omitted from consideration by ignoring the kinetic energy/momentum coordinate parts of S, E, A, and Cv. Hereon, we consider only the positional coordinate/potential energy part of these thermodynamic quantities.

Evaluation of the Heating Integral. The Clausius expression for the entropy change is given by eq 3. If the heat was absorbed at constant temperature T (as in a phase change such as ice melting), the entropy change would be

$$\Delta S|_T = \frac{\Delta E}{T} \quad (4)$$

The heat absorbed is just equal to ΔE, the internal energy change. Generally, T is not constant, and the temperature will rise from say T_1 to T_2 as the heat is absorbed, in a manner determined by the heat capacity. So the true entropy change satisfies

$$\frac{\Delta E}{T_2} < \Delta S < \frac{\Delta E}{T_1} \quad (5)$$

Equation 5 says that using the true energy change but assuming the temperature does not rise leads to an overestimate of the entropy change. Conversely, assuming that the temperature rises instantaneously to its final value leads to an underestimate

dx.doi.org/10.1021/ct300901x | J. Chem. Theory Comput. 2013, 9, 1164−1172

of the entropy change. Therefore, there is some as yet undetermined "Goldilocks'" temperature $T_1 < T_G < T_2$ that, when divided into the actual energy change, will give the true entropy change, namely

$$\Delta S = \int \frac{dq}{T} \equiv \frac{\Delta E}{T_G} \tag{6}$$

In fact, the existence of T_G is guaranteed the mean value theorem. Consider the case where the heat capacity Cv is constant over the range T_1:T_2. From the identity

$$\mathrm{Cv} = \frac{dE}{dT} = T\frac{dS}{dT} \tag{7}$$

the exact entropy change is given by

$$\Delta S = \mathrm{Cv}\ln\left(\frac{T_2}{T_1}\right) \tag{8}$$

Substituting eq 8 into eq 6, noting that $\mathrm{Cv} = \Delta E/(T_2 - T_1)$ and solving for T_G gives

$$T_G = \frac{T_2 - T_1}{\ln T_2 - \ln T_1} \tag{9}$$

For this case, T_G is actually a special type of average of T_1 and T_2, called here the van't Hoff average T_{vH}. It is straightforward to show that it has the required properties of an average, namely that $T_1 < T_{vH} < T_2$, and it is well-defined in the limit $T_2 \rightarrow T_1$ (see the Supporting Information).

Thus, the Clausius−van't Hoff (CvH) expression for entropy given by

$$\Delta S_{CvH} = \frac{\Delta E}{T_{vH}} = \frac{\Delta E(\ln T_2 - \ln T_1)}{T_2 - T_1} \tag{10}$$

is exact when Cv is constant over the temperature range T_1 to T_2. A continuous temperature integral has been replaced by a simple two point difference equation. The corresponding free energy change $\Delta A_{CvH} = \Delta E - T\Delta S_{CvH}$ follows immediately. Since ΔS is exact, both terms in the free energy are exact, so ΔA_{CvH} is also exact when Cv is constant.

If Cv is not constant, then eq 8 becomes

$$\Delta S = \int \frac{\mathrm{Cv}dT}{T} \tag{11}$$

To gauge the effect of variable heat capacity on the finite difference expression eq 10, let Cv vary linearly with temperature as

$$\mathrm{Cv}(T) = \mathrm{Cv}(T_1) + D(T - T_1) \tag{12}$$

From eq 11, the true entropy change is given by

$$\Delta S = \mathrm{Cv}(T_1)\ln(T_2/T_1) - DT_1\ln(T_2/T_1) + D\Delta T \tag{13}$$

where $\Delta T = (T_2 - T_1)$. Since $\Delta E = \int \mathrm{Cv}dT$, the mean energy change is now

$$\Delta E = \mathrm{Cv}(T_1)\Delta T - DT_1\Delta T + D(T_2^2 - T_1^2)/2 \tag{14}$$

The Clausius/van't Hoff estimate, using eq 14 for the true ΔE in eq 11, is

$$\Delta S_{CvH} \approx \mathrm{Cv}(T_1)\ln(T_2/T_1) - DT_1\ln(T_2/T_1) + D(T_2 + T_1)\ln(T_2/T_1)/2 \tag{15}$$

which differs from the exact entropy of eq 13 in the third term only. The ratio of the different third terms is

$$\frac{(T_2 + T_1)\ln(T_2/T_1)/2}{(T_2 - T_1)} = \frac{T_{av}}{T_{vH}} \tag{16}$$

which is the ratio of two different averages of T: the arithmetic average $T_{av} = (T_1 + T_2)/2$ and the van't Hoff average. This ratio is close to 1 even for large temperature increments. For example, doubling the temperature from 150 K to 300 K gives a ratio of averages of 1.04 (Table S1, Supporting Information). Alternatively, by taking the difference in entropy between eqs 13 and 15 and expanding the logarithm, the leading term in the error is

$$\delta S = \Delta S_{CvH} - \Delta S \approx \frac{D}{12}\left(\frac{\Delta T}{T_{av}}\right)^2 \tag{17}$$

Unless the temperature variation of heat capacity is very large, the difference between the exact Clausius entropy and its constant-Cv estimate will be small. Moreover, as eqs 16 and 17 indicate, the error is easily reduced by subdividing the heating into steps, so reducing $\Delta T/T_{av}$. The robustness of the Clausius−van't Hoff equation even for large Cv changes is illustrated for a model system in Figure 1. The model potential,

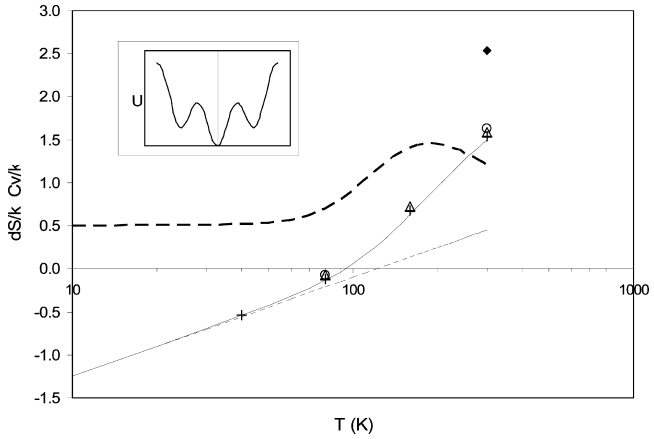


Figure 1. Entropy of a highly anharmonic potential $U = (1 - \cos(2x)) + 0.1x^2$ kcal/mol, x in Å (shape is depicted in inset). Heat capacity (−−), exact entropy by numerical integration of eq 1 (—), harmonic behavior ($S/k \propto 1/2 \ln T$) (...). Clausius van't Hoff entropy expression, eq 10, for entropy change from 10−300 K in one step (◆), two steps (○), three steps (Δ), or four steps (+).

$U = (1 - \cos(2x)) + 0.1x^2$, resembles a torsion potential used for amino acid side chains, and it was chosen because it has a very large anharmonicity: From 60 K to 300 K, the heat capacity increases by more than 250%. For comparison, in a previous study designed to test anharmonic corrections to the entropy, an anharmonic potential of the form $U = K(e^{-x} + x)$ was used.[32] This gives just a 20% increase in Cv over the same T range. The Clausius−van't Hoff difference equation was used to calculate the entropy change from 10 K to 300 K, a large range. Using just one step, the error, at 38%, is significant, as one might expect from such a drastic approximation to the quadrature. Breaking this into two steps (using T_{vH} to pick the intermediate T), the error drops to 5%. Interpolation of additional T points to produce three and then four step heating reduces the error to 3%.

dx.doi.org/10.1021/ct300901x | J. Chem. Theory Comput. 2013, 9, 1164−1172

The rather unexpected accuracy can be explained as follows: The Goldilocks temperature, defined in effect by the mean value theorem, is itself a form of average which cannot vary far from the average defined by T_{vH}, even for large changes in Cv. So T_{vH} is a good estimator of it. Consider the model anharmonic potential in the region of greatest anharmonicity, for example. Back calculation of T_G using the exact entropy change for the step T = 160 K to 300 K gives T_G = 221.6 K compared to T_{vH} = 222.7 K, a difference of just 1.1 K.

In summary, one can integrate over a large temperature range with a rather small number of steps. Since T is given (in Monte Carlo methods) or tightly controlled (in well equilibrated molecular dynamics simulations), the principal computational hurdle is accurate evaluation of a single average system property at each temperature, namely E.

Component Analysis of the Entropy. Analysis of the different contributions to the entropy can in principle provide deeper insight into the behavior of the molecules. A natural way to do this is to consider the contribution from each of the N degrees of freedom independently (the first order term), the contribution from $N!/(2!(N-2)!)$ second order (pairwise correlation) terms, then three-body correlations, and so on, as in the mutual information expansion of eq 2. Carrying this through the full expansion, a total of $2^N - 1$ terms result. Even one of the simpler entropy models, the quasi-harmonic model (where there are only first order and second order terms given by the diagonal and off-diagonal terms of the covariance matrix, respectively) has $O(N^2)$ contributions. With this kind of decomposition, the large number of terms makes it difficult to extract physical insights. The Clausius–van't Hoff approach provides an alternative and simpler decomposition of the entropy. If the Hamiltonian is the sum of m terms

$$U = U_1 + U_2 + ...U_m \tag{18}$$

then the mean potential energy is immediately decomposable into the corresponding m terms

$$E = \langle U \rangle = \langle U_1 \rangle + \langle U_2 \rangle + ... \langle U_m \rangle = E_1 + E_2 + ...E_m \tag{19}$$

Since the positional contribution to entropy is given by $\Delta S = \int (dE/T)$, then the corresponding decomposition is

$$\Delta S = \int_o^T \frac{dE_1}{T} + \int_o^T \frac{dE_2}{T} ... \int_o^T \frac{dE_m}{T} = \Delta S_1 + \Delta S_2 + ...\Delta S_m \tag{20}$$

This follows directly from Clausius' expression for the entropy. Its validity does not depend on the Clausius–van't Hoff approximation or other approximations used for numerical evaluation. This form of decomposition is analogous to that previously demonstrated for the Helmholtz free energy A.[33] There, a temperature integration of the form $\int_0^\beta ...d\beta$, $\beta = 1/kT$ provided a decomposition of A into the analogous m terms. In the Clausius–van't Hoff approximation, the entropy components are simply

$$\Delta S_i = \frac{\Delta E_i}{T_G} \approx \frac{\Delta E_i}{T_{CvH}},\ i = 1, m \tag{21}$$

with $O(N)$ terms. The corresponding breakdowns of E, and thus A, follow. The Hamiltonian may be broken down in various ways. For example, for a solute in water, an informative decomposition is into solute–solute, solute–water, and water–water entropy and free energy terms. These may be further broken down into group or residue contributions. An alternative form of decomposition is into types of energy terms, such as van der Waals and electrostatics. Although this is harder to interpret in a physical sense, it is similar in spirit to decompositions used previously to analyze free energy changes in free energy perturbation studies.[34,35]

In the decomposition based on the Clausius equation, one may ask where the correlation terms of eq 2 have gone. The answer is exactly that found previously for the decomposition of A:[33] thermal integration partitions of each of the correlation terms among the first order terms of the coordinates it involves. So a correlation term between say three coordinates (ijk) is partitioned into each of the three single coordinate entropy terms i, j, and k. Specifically, in our previous free energy analysis, the coupling terms appear as terms in the cumulant expansion of A, and each such term is partitioned among its first order terms $i, j...k$ in proportion to the exponents of the ith, jth...kth coordinates in the cumulant.[33]

Non-*T*-Dependent Entropy Changes. So far, only an entropy change due to a change in temperature has been considered. A more common need is for the change in entropy of a molecule due to a change in its environment or state. Examples are the change in conformational entropy of two molecules upon binding and the change in entropy of a protein upon folding. To apply the Clausius equation to this situation, one would calculate the absolute entropy for state I by integrating the heat flow from 0 K to T:[34]

$$S_I(T) = \int_0^T \frac{dq^I}{T} = \int_0^T \frac{C_V^I(T)dT}{T} \tag{22}$$

The process is similar for state II. The entropy change for process I→II is then the difference in the two third-law entropies $\Delta S(T) = S_{II}(T) - S_I(T)$. To make this practical, the CvH approximation would again be used to replace the continuous integration of eq 22 with a relatively small number of heating steps. [Note that for the initial T step from 0 K, T_{vH} is not defined. One could simply approximate T_G by T_{AV} or just take a small initial T step and neglect its entropy. Since Cv ≈ 0 for this step, negligible error in the entropy will be introduced.] Unfortunately, the classical treatment of the Hamiltonian currently used in most macromolecular simulations leads to a problem: The heat capacity does not approach zero as $T \to 0$. This is illustrated in Figure 2 for the classical treatment of a Hamiltonian consisting of a single quadratic (harmonic) term. Here, Cv = $k/2$, and so from eq 22 the entropy decreases

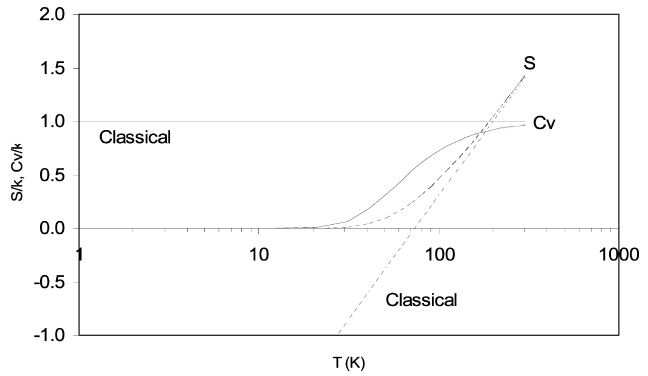


Figure 2. Temperature dependence of Cv (solid line) and S (dashed line) for a harmonic potential: Quantum and classical approximation lines are shown.

dx.doi.org/10.1021/ct300901x | J. Chem. Theory Comput. 2013, 9, 1164–1172

without a limit as $T \rightarrow 0$. The proper quantum behavior is also shown on the figure. Here Cv→0, so $S \rightarrow 0$.

The actual Hamiltonians used for biomolecules are considerably more complex than a simple harmonic potential, but one can see, by expanding the potential around the minimum in a Taylor series and setting $\nabla U = 0$, that the leading term is harmonic and the same unphysical heat capacity behavior results from the classical treatment. Expressing the problem in terms of eq 1, the U-shape of the Hamiltonian allows, in a classical treatment, the probability distribution to narrow without limit as the system is cooled, so $S \rightarrow -\infty$. The ideal solution would be to use a fully quantum treatment, but this is not practical for most biological systems. Another approach would be to modify the Hamiltonians in some way so that they have physically realistic heat capacity behavior at low temperatures in a classical treatment. In lieu of such a force field development, two other strategies based on temperature quenching are proposed.

Protocol 1. The system is cooled in steps to a low temperature T_q until the system behaves harmonically. The CvH method is used to calculate the entropy change during this cooling. Whether or not the system behaves harmonically during the initial cooling steps is immaterial to the CvH method. Results from the anharmonic potential example of Figure 1 indicate that a small number of steps will suffice to quench the anharmonicity. In any case, a simple internal check that harmonicity is reached is that between the ith and $(i + 1)$th steps of cooling $\Delta S_{\mathrm{CvH}}(i \rightarrow i+1) = {}^1/_2 nk \ln(T_i/T_{i+1})$, where n is the number of degrees of freedom. The onset of the harmonic regime is shown in Figure 1 for $T < 60$ K. Once the system behaves harmonically, any suitable method such as the quasiharmonic method[10] can be used to calculate the entropy at T_q accurately. The two contributions S_{harm} and ΔS_{CvH} are then added. To obtain the entropy change for process I→II, this protocol is applied to the molecule(s) in states I and II and the difference taken. If the free energy difference is required, ΔE is obtained from the difference in mean energies in states I and II at the ambient temperature, and then $\Delta A = \Delta E - T\Delta S$.

Protocol 2. The system is minimized (so effectively $T = 0$ K) and the normal-mode frequencies ν_j are obtained by diagonalization of the Hessian matrix.[9] The entropy of the system at T is calculated from the usual semiclassical expression

$$S_{nm}(T) = \sum_j \ln\left(\frac{kTe}{h\nu_j}\right) \tag{23}$$

(here $S_{nm}(T)$ does include both kinetic and potential terms). $S_{nm}(T)$ would be the true entropy if the system was thermalized at temperature T but retained its harmonicity. Using the CvH method, the change in excess entropy (relative to that of a harmonic system with the frequencies of eq 23) from heating from 0 K to T is calculated using a suitable number of heating steps as follows: If the system remained harmonic, the change in potential energy from a temperature increment ΔT would be $nk\Delta T/2$, and the corresponding contribution to the entropy change would be $nk\Delta T/(2T_{\mathrm{vH}})$. The change in excess potential energy is thus $\Delta E_i^{\mathrm{excess}} = \Delta E_i - nk\Delta T_i/2$, where ΔE_i is the change in mean potential energy of the system going from T_i to T_{i+1}(the kinetic energy always goes as $nk\Delta T_i/2$ so it has no "excess" relative to a harmonic system). Then

$$\Delta S_{\mathrm{CvH}}^{\mathrm{excess}}(i \rightarrow i+1) = \Delta E_i^{\mathrm{excess}}(\ln T_{i+1} - \ln T_i)/(T_{i+1} - T_i) \tag{24}$$

The total entropy is

$$S_{nm}(T) + \sum_i \Delta S_{\mathrm{CvH}}^{\mathrm{excess}}(i \rightarrow i+1) \tag{25}$$

In practical terms, protocol 1 has the advantage that one can calculate the entropy at any temperature T_q in the harmonic regime and combine this with the entropy of quenching from T to T_q—the calculated entropy for the system at T from eq 25 should be the same. In a simulation of finite extent and accuracy, one may combine results obtained from quenching to several different temperatures in order to improve accuracy and obtain an estimate of the numerical uncertainty. Conversely, in protocol 2, calculation of the harmonic part of the entropy (which is based on eq 1 and is the "difficult" part to sample relative to the S_{CvH} part due to the large number of terms) is dependent on the accuracy of a single T point, the minimized ($T = 0$ K) state.

Protocol 1 is framed in terms of quenching. However, one may equally well perform the steps in reverse and heat the system. In fact, many MD simulation protocols start by minimizing a structure and then heating in controlled stages to avoid distorting the system with an uneven distribution of kinetic energy. This is usually viewed as a necessary evil to be done as expeditiously as possible. To the contrary, this heating, if done in such a way that reliable average E values are obtained, could be used to extract useful entropy estimates via the Clausius−van't Hoff equation. A final point is that in the CvH method, each T point is an independent, equilibrium simulation. If it turns out in retrospect that a particular T step is too large, an additional T point can be inserted, the neighboring ΔE and T_{vH} values recalculated, and hence an improved ΔS_{CvH} recalculated with no loss of effort from the earlier simulations.

Molecular Dynamics. Molecular dynamics (MD) simulations were performed using the CHARMM molecular modeling package[35,36] using the CHARMM27 force field.[36] Additional simulation parameters and conditions were chosen as in our previous study of ligand entropy.[37] For ligands not contained in the CHARMM27 force field, bond and van der Waals parameters were transferred from chemically homologous groups or moieties. Atomic charges were obtained using the electronegativity neutralization method, Qequil.[38] After the assignment of parameters and building of hydrogens, the structures were exhaustively minimized using the ABNR minimizer. Molecular dynamics simulations were performed with a time step of 1 fs and a nonbonded cutoff of 14 Å using the shift method. A distant-dependent dielectric of $4r$ was used as an implicit solvent model. While not particularly realistic, the focus of this study is the methodology of entropy extraction from simulations. Simulations are sufficiently rapid to ensure that configurational entropy estimates converge in a reasonable amount of time. An implicit solvent model also has the advantage in the Clausius method in that it can be kept T-invariant. Thus, the only explicit degrees of freedom that can give rise to an entropy *change* are internal conformational ones. Simulations were run with Langevin dynamics, using an atomic friction coefficient of 70 ps^{-1} scaled by the fractional surface accessibility of the atom. The Langevin dynamics approach was selected because it models solvent fluctuation effects missing in

an implicit solvent model, and it improves conformational sampling. It also provides excellent temperature control. Simulations were equilibrated for 1 ns and sampled for 100 ns. Frames were saved every 0.1 ps for entropy calculations.

Calculation of Ligand Configurational Entropy from Coordinate pdf's. For each coordinate frame molecular coordinates were converted to a nonredundant internal coordinate representation of torsion angles, bond angles, and bond lengths. For ligands, one-dimensional pdf's of the internal coordinates were accumulated using a bin size of 3° for torsions, 1° for bond angles, and 0.01 Å for bond lengths. The uncorrelated entropy estimate (i.e., first order MIE) was obtained by

$$S^{1D} = -k \sum_i \sum_m p_m^i \ln p_m^i \tag{26}$$

where k is the Boltzmann constant and $p_m{}^i$ is the probability of the ith internal coordinate being in the mth bin.

The coordinate fluctuation covariance matrix, whose elements are given by $\sigma_{ij}{}^2 = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle$ was calculated from the internal coordinate frames, where $\langle ... \rangle$ indicates an average over the trajectory. The fluctuation correlation matrix was obtained from the coordinate fluctuation covariance matrix as $C_{ij} = \sigma_{ij}{}^2/(\sigma_{ii}{}^2\sigma_{jj}{}^2)^{1/2}$. The entropy in the quasi harmonic model is given by[10]

$$S^{qh}/k = \frac{1}{2} \ln((2\pi e)^n |\sigma_{ij}^2|) \tag{27}$$

The entropy in the Boltzmann−quasiharmonic model of Di Nola et al. is given by

$$S^{BQH} = S^{1D} + \frac{1}{2} k \ln|C_{ij}| \tag{28}$$

where the second term represents the correction to the first order MIE due to correlations at the quasiharmonic level of approximation. Householders elimination was used to calculate the determinants of σ_{ij} and C_{ij}.[39]

Monte Carlo Simulations of Water. A box of 216 water molecules was simulated at constant volume and constant temperature using the TIP4P potential, Monte Carlo sampling, and minimum image periodic boundary conditions. No nonbonded cutoff was used. One million steps of equilibration were performed, followed by 10 million steps of sampling. Simulations were run at 300 K, and at several decreasing temperatures down to 5 K. The quenching entropy relative to T = 300 K was calculated at each temperature using the CvH eq 10 until the system demonstrated harmonic behavior. Then, the quasiharmonic method was used to calculate the entropy of the quenched state in the harmonic regime using eq 27. For the rigid body TIP4P molecule, the coordinate system used was the position of each water center (O atom) in Cartesian coordinates and its orientation expressed in Cartesian angles α, $\sin(\beta)$, γ of the local molecular frame relative to the water box coordinate axes resulting in a $(6 \times 216)^2$ element σ_{ij} matrix. The vaporization entropy per mole of water at 300 K was then obtained as

$$S_{vap} = S_{quench}/n_w - (S^{qh}/n_w - S_{ref}) \tag{29}$$

where n_w is the number of water molecules and S_{ref} is the gas phase reference state entropy for a rigid TIP4P water molecule given by

$$S_{ref} = k \ln(8\pi^2 V_{mol}) \tag{30}$$

evaluated at the simulation density of V_{mol} = 28.3 Å^3.

RESULTS

Benchmarking Methods for Calculating Entropy. The numerical accuracy of the Clausius/van't Hoff equation depends principally on accurate evaluation of a single average quantity, E. This should be contrasted with evaluation of S using eq 1. Here, numerous coordinate pdf's need to be individually estimated and then their entropy contributions summed. At some point, higher order pdf's must simply be neglected in the evaluation of eq 1. The CvH method, in contrast, implicitly includes such correlation effects to all orders. Some methods also make simplifying assumptions about the form of the pdf's, for example, quasiharmonic methods. One use for the Clausius/van't Hoff method then is to assess the error introduced by these and other approximations. This is illustrated by examining the accuracy of the Boltzmann−quasiharmonic (BQH) method of DiNola et al.[1] (eq 28). This is an approximation to eq 1 combining the uncorrelated one-coordinate entropy terms with a second order correlation correction term treated at the quasi-harmonic level. Since the BQH method is only second order in N, it is computationally tractable. Moreover, the second order term requires only coordinate variances and covariances, rather than full 2-D pdf's, further reducing memory and sampling requirements. Thus, the BQH method is tractable even for large systems such as an entire protein.[1,37] Five ligands ranging in size from ∼10 to ∼100 atoms (Table 1) were used in this analysis. The smallest, alanine dipeptide is a widely used test peptide for simulation. The other four are ligands of proteins, namely carboxyl arabinotyl 2,5 phosphate, a ligand of RUBISCO; biotin, a ligand of streptavidin; digoxin, a ligand of antibody 26−10; and peptide PEPTAPPEE, a ligand of HIV-1 protease. The conformational entropy change was obtained as described previously.[37] Molecular dynamics (MD) simulations of each molecule were run at constant temperatures of T = 280 K and T = 320 K using an implicit solvent model. The Clausius−van't Hoff (CvH) entropy changes were computed from the difference in mean potential energy using eq 10. Internal coordinate (bond length, angle, and torsion) pdf's and covariance fluctuation matrices were used for the calculation of the corresponding BQH entropy change. Entropy estimates calculated from the CvH method converged 10 or more times faster than those from the BQH method (Supporting

Table 1. Solute Conformational Entropy: Comparison of Methods[a]

compound[b]	atoms	method[c]			
		CvH	BQH	1D	QH
alanine dipeptide	22	7.9	7.9	7.2	7.8
carboxyl arabinotyl 2, 5 phosphate (1RBO)	30	12.5	13.5	13.2	14.9
biotin (1STP)	31	51.7	63.0	67.0	67.1
digoxin (1IGI)	62	11.8	15.1	15.3	23.8
peptide PEPTAPPEE (1M4P)	128	24.6	26.6	27.5	31.6

[a]Entropy change in cal/mol/K for T_2 = 320 K vs T_1 = 280 K. [b]PDB entry code used for starting structure is indicated in parentheses. [c]CvH: Clausius van't Hoff. BQH: Boltzmann-quasiharmonic. 1D: using one-dimensional pdf's only. QH: quasiharmonic.

Information Figure S1), so simulation length was determined by the time needed for the BQH method to converge. A comparison of the entropy changes calculated from the two methods for the five solutes is shown in Figure 3. As in our

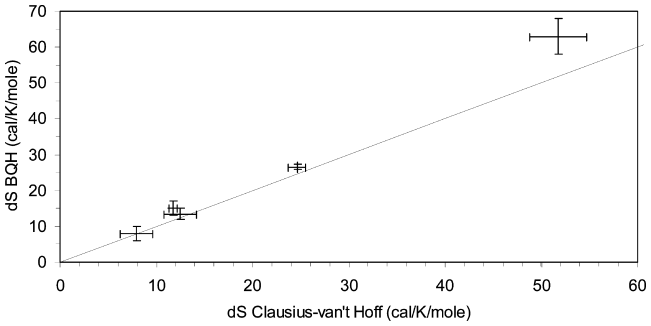


Figure 3. Comparison of the Boltzmann–quasiharmonic method of DiNola et al.[1] eq 28, with the Clausius–van't Hoff method, eq 10. Error bars are obtained from the variance in batch means.

previous study, the agreement of BQH with CvH is very good for most of the ligands, considering the substantial approximations made in the BQH method. However, the BQH method gives systematically larger entropies. One can attribute this to the approximations made in the BQH method, namely, neglect of third-order and higher correlations and/or use of the quasi-harmonic treatment of the second order correlation term. The latter is expected, *a priori*, to cause systematic overestimation of S. Entropies were also calculated using just the 1D pdf's (S_{1D}), i.e., neglecting all correlations (eq 26), and then assuming that all the pdf's were Gaussian—the full quasiharmonic model of eq 27 (S_{QH}). The relative accuracy of the three methods is expected to be $S_{QH} \approx S_{1D} < S_{BQH}$ as the pdf treatment is improved. As Table 1 shows, this expectation is borne out, the points moving toward the "exact" S_{CvH} values. Alanine dipeptide is small enough that S_{QH} is already quite accurate, and all three approximations to eq 1 give comparable accuracy.

Benchmarking can even be applied retrospectively to entropy estimates in the literature, providing the minimal information of E and T needed by the CvH method is available. For example, Killian et al.[21] compared three different approximations to eq 1, including MIE, for evaluating the conformational entropy of dichloroethane at three different temperatures 300 K, 500 K, and 1000 K. They also tabulated the mean potential energies. From this, one can calculate the entropy change using the Clausius–van't Hoff method (Table 2). The agreement is almost exact throughout, and the precision is good enough that one can verify the expected improvement in accuracy as the order of the mutual information expansion is increased: The difference relative to Clausius–van't Hoff drops from about 1.5% with MIE1 to about 0.5% with MIE2.

Diffusive Systems. It is well-known that the narrower and more monodisperse the pdf distributions are, the easier the evaluation of eq 1 is. The quasi-harmonic method, and methods based on it, in fact assume a single peaked, multivariate Gaussian distribution. For these reasons, many methods based on evaluation of eq 1 are best suited to oscillatory rather than diffusive motions, such as a protein undergoing excursions around a well-defined folded structure. Since the CvH approach requires only mean energies, and not pdf's, it is not limited in this way. To demonstrate this, and also to illustrate the

Table 2. Benchmarking Entropy Calculations[a]

	M2 entropy[b]	MIE 1 entropy	MIE 2 entropy	⟨E⟩	CvH entropy[c]
S at 300 K	−21.10	−18.60	−21.00	5.72	
ΔS at 500 K[d]	9.72	9.86	9.72	9.51	9.68
ΔS at 1000 K[d]	22.46	22.63	22.41	18.68	22.29

[a]Data from Killian et al.[21] for dichloroethane. Entropies in cal/mol/K, energies in kcal/mol. [b]M2: Second generation mining minima. MIE1: Mutual information expansion to second order. MIE2: Mutual information expansion to third order. [c]Entropy relative to that at 300 K using eq 10 and the tabulated mean potential energies. [d]Absolute entropy at 300 K and entropy relative to that at 300 K for the other two temperatures.

application of protocol 1 above, the method was used to calculate the vaporization entropy of liquid water. The results are shown in Figure 4. The CvH entropy decreases smoothly,

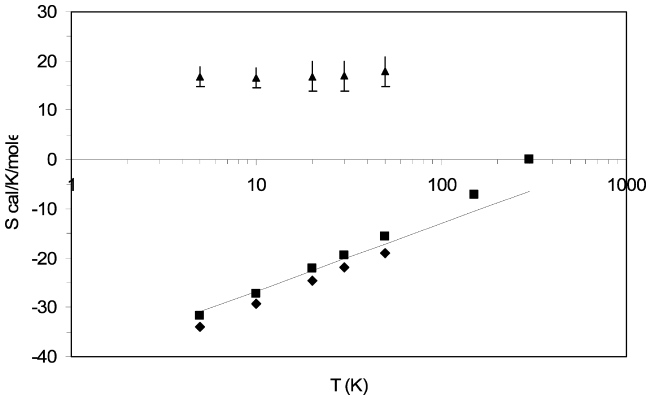


Figure 4. Water vaporization entropy from Monte Carlo simulation of TIP4P waters. (■) Entropy of quenching from 300 K, calculated using the Clausius–van't Hoff equation. Harmonic behavior is indicated by the line with a slope of $^1/_2 n \ln(T)$. (◆) Entropy of system in harmonic regime, calculated from the coordinate fluctuation covariance matrix using the quasi-harmonic equation. (▲) Entropy of vaporization at 300 K calculated via quenching to the indicated temperature.

and the onset of harmonicity is seen by the approach to the $^1/_2 nk \ln(T)$ line. In the anharmonic regime, the entropy is in positive excess relative to the harmonic behavior. At the lower temperatures, the water molecules are frozen and make smaller and smaller oscillations around fixed positions, leading to harmonic behavior. As expected from the theoretical analysis, large T steps can be taken so the harmonic regime can be achieved in a very small number of steps. The quasiharmonic entropy for T < 50 K also follows a line of slope $^1/_2 nk \ln(T)$, a self-consistent indication that the harmonic regime has been reached.

Quenching from 300 K to successively lower temperatures results in a larger decrease in entropy. However, the quasiharmonic entropy in the cooler quenched state is also lower, so the increase in entropy upon transfer to the vapor reference state is correspondingly larger. The vaporization entropy is the sum of these two contributions. Thus, in the harmonic regime where the quasiharmonic model should be accurate using any of the five final quench points from 5 K to 50 K, gives the same vaporization entropy at 300 K to within numerical error. Seven simulation temperatures were used in the figure to illustrate the onset of harmonicity and the

invariance of the vaporization entropy once harmonicity is reached. This is more than would be needed in practice, and as few as 3−4 T steps would be enough to obtain the same accuracy of vaporization entropy. The average value of 15 ± 1 cal/K/mol is in good agreement with the previous simulation value of 13.1 cal/mol/K obtained by Jorgensen et al. using the BOSS Monte Carlo program[40,41] but somewhat larger, which we attribute to small differences in simulation details and sampling variation. The result is in agreement with the experimental value of 14.0 cal/mol/K to within estimated simulation precision limits.[42]

DISCUSSION

The Clausius expression for entropy provides an alternative to the Boltzmann expression for the calculation of entropy changes, via temperature integration rather than accumulation of coordinate pdf's. Analytical and numerical error analysis shows that the integration can be done accurately using a small number of large temperature steps involving the calculation of one average system property, the mean internal energy. The key to this accuracy is robust estimation of the mean value theorem temperature over the heating step by a form of thermodynamic average temperature, introduced here as the van't Hoff average. This makes the method computationally practical for systems with many degrees of freedom, such as biological molecules. One immediate application is for benchmarking other methods for entropy calculation, such as statistical mechanical ones based on the evaluation of coordinate probability distribution functions. The benchmarking can be done on the actual systems one is interested in studying, rather than simple model systems. This benchmarking is used to provide a quantitative assessment of one such statistical mechanical method, the Boltzmann−quasiharmonic (BQH) method.[1] The BQH method is one of considerable potential as it is computationally tractable for systems as large as peptides and whole proteins.

The Clausius expression also provides an alternative way of decomposing the entropy into components. This decomposition is linear in the number system components, rather than exponential. This is physically appealing in the interpretation of entropy changes in systems with a large number of degrees of freedom.

Implicit solvent models are widely used in macromolecular simulations. They combine the accurate description of solvation effects with the rapidity needed to get well converged sampling of solute conformations. Thus, many recent studies of molecular conformational entropy calculations have been performed with implicit solvent.[15,21,23,39] The CvH estimate of entropy converges particularly rapidly for implicit solvent simulations (Supporting Information Figure S1), more than 10 times faster than the BQH method using the same simulation snapshots, requiring evaluation of coordinate pdf's. Preliminary studies with explicit solvent (Supporting Information Figure S2) indicates the CvH method converges as rapidly with *explicit* solvent as the BQH method does with *implicit* solvent. This indicates the feasibility of the CvH method for sub-microsecond simulations.

As described here and in our previous study of solute conformational entropy,[37] one can also take take advantage of the implicit solvent treatment in the heating integration by "freezing" the implicit solvent term. The only temperature dependence then comes from the change in conformational entropy of the solute, thus isolating the contribution of interest. In effect, this is just an application of the entropy decomposition described by eq 21.

A barrier to the wider use of the Clausius expression for entropy evaluation is the difficulty of evaluating the absolute or third law entropy by integration from $T = 0$ K. This is because with a classical treatment, biomolecular simulations have unphysical behavior at low temperatures; namely, the heat capacity does not go to zero as $T \rightarrow 0$. Several solutions are proposed here. Two protocols combine the Clausius method with harmonic or quasiharmonic methods. The Clausius part takes care of the entropy changes at higher temperatures where the molecular motions are highly anharmonic and correlated. This is the region where entropy estimation via probability distributions is difficult. Application of one of these protocols to the calculation of liquid water vaporization entropy is described. This also illustrates the utility of the Clausius method for entropy estimation where motions are diffusive, rather than oscillatory.

A potential difficulty with the quenching method is that as T is lowered, sampling of configurational space becomes slower. The system might settle into a local minimum far from the global minimum and appear prematurely converged. This difficulty was not apparent in the application here to estimation of pure water entropy, where good results were obtained. However, it is more likely to occur in larger biomolecular systems where there could be many local minima. Sampling difficulties are unavoidable with any large system, and quenching would tend to exacerbate them. On the other hand, existing strategies, such as multiple independent trajectories, and replica exchange may well be applicable here too.

Another solution to the absolute entropy problem is to modify Hamiltonians so that they have more physical behavior at low temperatures in the classical treatment. In a fully quantum treatment, as $T \rightarrow 0$, the system drops into a single ground state. From the original Boltzmann definition of entropy, $S = k \ln \omega$, where ω is the number of microstates, so $S = 0$, in agreement with the third law. Analysis of the origin of the unphysicality in the classical treatment shows that there is no counterpart of dropping into a single state: With a U-shaped potential around the minimum, the pdf can narrow without limit. However, by analogy with the single quantum state behavior, if the potential is "flat-bottomed," one can show that in a classical treatment $Cv \rightarrow 0$ as $T \rightarrow 0$; i.e., the system behaves as though it is dropping into a single "state" defined as the floor of the potential. Hence, S goes to a fixed limit, enabling the third law to be applied. While specific implementation in this direction is outside the scope of this work, it is suggested that this or other strategies would be a fruitful way to develop a classical treatment with more realistic thermodynamic properties.

A final incentive for future development and application of the Clausius approach to entropy calculations comes from the increased use of multiscale and hybrid simulations of macromolecules. They allow a trade-off between system size, computational tractability, and accuracy. One example is the explicit solute, implicit solvent model discussed above. Other examples include combined quantum mechanics/molecular mechanics (QM/MM) simulations often used for enzymes, and coarse grained representations of lipids in membrane simulations. Different parts of the system may be represented with different physical models, or levels of detail, some degrees of freedom may only be implicit. Determining the entropy in

such systems via evaluation of pdf's is challenging because a consistent and nonredundant set of "coordinates" must be used. Considering QM/MM, for example, how does one combine the QM wave function representation with the MM nuclear positions, particularly since atoms at the boundary between different representations may have dual character? The Clausius method bypasses such difficulties since it requires only the mean energy and temperature. These are presumably well-defined and unambiguous in any physically meaningful hybrid model.

■ ASSOCIATED CONTENT

Supporting Information

Contains proof that the van't Hoff expression is a true average of two temperatures, a comparison of van't Hoff and arithmetic averages, and plots showing the convergence behavior of the CvH and BQH methods for entropy calculation. This information is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 215-573-3506. E-mail: sharpk@mail.med.upenn.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

I would like to thank Art Palmer for helpful discussion on the Clausius−van't Hoff error analysis. Support is acknowledged from the National Resource for Biomedical Supercomputing and the Pittsburgh Supercomputing Center through NIH Award RC2GM093307 to CMU through the NRBSC.

■ REFERENCES

(1) Di Nola, A.; Berendsen, H.; Edholm, O. *Mol. Phys.* **1984**, *17*, 2044−2050.
(2) Monod, J.; Wyman, J.; Changeux, J.-P. *J. Mol. Biol.* **1965**, *12*, 88−118.
(3) Cooper, A.; Dryden, D. T. F. *Eur. Biophys. J.* **1984**, *11*, 103−109.
(4) Marlow, M. S.; Dogan, J.; Frederick, K. K.; Valentine, K. G.; Wand, A. J. *Nat. Chem. Biol.* **2010**, *6*, 352−358.
(5) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325−329.
(6) DeLorbe, J. E.; Clements, J. H.; Teresk, M. G.; Benfield, A. P.; Plake, H. R.; Millspaugh, L. E.; Martin, S. F. *J. Am. Chem. Soc.* **2009**, *131*, 16758−16770.
(7) Freire, E. *Drug Discovery Today* **2008**, *13*, 869−874.
(8) Boltzmann, L. *Weiner Ber.* **1877**, *76*, 373−435.
(9) Go, N.; Scheraga, H. *J. Chem. Phys.* **1969**, *51*, 4751−4767.
(10) Karplus, M.; Kushick, J. *Macromolecules* **1981**, *14*, 325−332.
(11) Levy, R.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370−1374.
(12) Case, D. A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 285−290.
(13) Edholm, O.; Berendsen, H. *Mol. Phys.* **1984**, *51*, 1011−1028.
(14) Lazaridis, T.; Paulaitis, M. *J. Phys. Chem.* **1992**, *96*, 3847−3855.
(15) Head, M. S.; Given, J. A.; Gilson, M. K. *J. Phys. Chem. A.* **1997**, *101*, 1609−1618.
(16) Luo, H.; Sharp, K. A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 10399−10404.
(17) Baron, R.; Van Gunsteren, W. F.; Hünenberger, P. H. *Trends Phys. Chem.* **2006**, *11*, 87−122.
(18) Cheluvaraja, S.; Meirovitch, H. *J. Chem. Phys.* **2006**, *125*, 024905−024913.
(19) Wang, J.; Bruschweiler, R. *J. Chem. Theory Comput.* **2006**, *2*, 18−24.
(20) Hnizdo, V.; Darian, E.; Fedorowicz, A.; Demchuk, E.; Li, S.; Singh, H. *J. Comput. Chem.* **2007**, *28*, 655−668.
(21) Killian, B. J.; Kravitz, J.; Gilson, M. K. *J. Chem. Phys.* **2007**, *127*, 024107.
(22) Hnizdo, V.; Tan, J.; Killian, B. J.; Gilson, M. K. *J. Comput. Chem.* **2008**, *29*, 1605−1614.
(23) Somani, S.; Killian, B. J.; Gilson, M. K. *J. Chem. Phys.* **2009**, *130*, 134102.
(24) Li, D.-W.; Showalter, S.; Bruschweiler, R. *J. Phys. Chem. B* **2010**, *114*, 16036−16044.
(25) Singh, N.; Warshel, A. *Proteins* **2010**, *78*, 1724−1735.
(26) Shannon, C. E. *Bell Syst. Tech. J.* **1948**, *27*, 379−423.
(27) Jaynes, E. T. *Phys. Rev.* **1957**, *106*, 620−630.
(28) Matsuda, H. *Phys. Rev. E* **2000**, *62*, 3096−3102.
(29) King, B. M.; Silver, N.; Tidor, B. *J. Phys. Chem. B* **2012**, *116*, 2891−2904.
(30) Paninski, L. *Neural Comput.* **2003**, *15*, 1191−1253.
(31) Clausius, R. *Ann. Phys. (Berlin, Ger.)* **1854**, *93*, 481.
(32) Rojas, O.; Levy, R. M.; Szabo, A. *J. Chem. Phys.* **1986**, *85*, 1037−1043.
(33) Brady, G. P.; Szabo, A.; Sharp, K. A. *J. Mol. Biol.* **1996**, *263*, 123−125.
(34) Dugdale, J. S. *Entropy and Low Temperature Physics*; Hutchison: London, 1966.
(35) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(36) MacKerell, A. D.; Brooks, B.; Brooks, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and Its Parameterization with an Overview of the Program. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley& Sons: Chichester, U. K., 1998; Vol. *1*; pp 271−277.
(37) Harpole, K.; Sharp, K. *J. Phys. Chem. B* **2011**, *115*, 9461−9472.
(38) Yang, Q.; Sharp, K. A. *J. Chem. Theory Comput.* **2006**, *2*, 1152−1167.
(39) Press, W.; Flannery, B.; Teukolsky, S.; Vetterling, W. *Numerical Recipes*; Cambridge University Press: Cambridge, U. K., 1986.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(41) Jorgensen, W.; Blake, J.; Buckner, J. *Chem. Phys.* **1989**, *129*, 193−200.
(42) Chemistry WebBook, Vol. *2012*. http://webbook.nist.gov/chemistry/ (accessed Dec. 2012).