Free energy basin-hopping

K.H. Sutherland-Cash [a], D.J. Wales [a,*], D. Chakrabarti [b]

[a] University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, UK
[b] School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

ABSTRACT

A global optimisation scheme is presented using basin-hopping with the acceptance criterion based on approximate free energy for the corresponding local minima of the potential energy. The method is illustrated for atomic and colloidal clusters and peptides to examine how the predicted global free energy minimum changes with temperature. Using estimates for the local free energies based on harmonic vibrational densities of states provides a computationally effective framework for predicting trends in structure at finite temperature. The resulting scheme represents a powerful tool for exploration of energy landscapes throughout molecular science.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Global optimisation is an important tool for structure prediction throughout molecular science, as well as for soft matter and condensed matter systems. In particular, numerous methods have been described that aim to identify the global minimum on the potential energy surface (PES) [1–4]. Although the PES of a many-body system is high dimensional and often complex, searches on the PES can usually be based on accurate evaluation of the potential energy function. In contrast, if the temperature is high enough for entropy to play a role, then further sampling may be necessary. Weak interactions that correlate with low frequency vibrational modes can be important here, suggesting that a systematic approach to investigate such effects could be very useful. However, we do not wish to incur the computational expense of constructing entire free energy landscapes based upon projections onto chosen order parameters, or sample extensively in a more global fashion [5–11]. Instead we prefer to focus on the local free energy associated with particular minima on the potential energy surface, and investigate how the competition between alternative structures changes with temperature. We do not propose to compute the free energy for states associated with sets of potential energy minima. Such states would be important for analysis of phase transitions, where the high entropy phase is stabilised by landscape entropy, i.e. the configuration space associated with multiple minima that lie relatively high in potential energy. Rather, we consider the regime where structure is still well defined in terms of competition between configurations associated with distinct potential energy minima. Cross-overs

in stability with temperature are then associated with the local well entropies, which will increase more rapidly with temperature for structures that have lower vibrational frequencies. In fact, we could also lump such minima, for example, using a free energy barrier threshold [12]. This sort of procedure has been used to construct free energy disconnectivity graphs [10,13], but it requires additional information to define connectivity.

In the present contribution we present a methodology that provides an estimate of free energy of the local minima on the PES on the fly as they are visited during global optimisation. To this end, we adapt the basin-hopping (BH) global optimisation method [2,14,15], which has been used for a wide variety of molecular and soft and condensed matter, to explore the *free* energy landscape, as opposed to the *potential* energy landscape. Our approach relies upon the superposition approach [16–21] to thermodynamics, which expresses the global partition function, $Z(T)$, in terms of contributions from the catchment basin of each local minimum. In the canonical ensemble, which is the focus of this letter, we have

$$Z(T) = \sum_i Z_i(T), \tag{1}$$

where $Z_i(T)$ is the partition function of minimum i at temperature T. Z_i thus only draws contributions from the catchment basin [22,18] for which steepest-descent paths converge to minimum i. Here we will employ the harmonic normal mode approximation to obtain the vibrational density of states for each minimum. This harmonic superposition approach has been used successfully for a variety of applications in previous work [21,23–26]. Various schemes to treat anharmonic corrections have been proposed [17,20,25,27–30], but we anticipate that the FEBH approach will generally be applied using harmonic vibrational densities of states

in the first instance. Once such a survey has been conducted, it would certainly be possible to admit vibrational anharmonicity, and the procedure that employs a model anharmonic form used in the basin-sampling approach to thermodynamics [25] could provide an attractive framework. We will report on such extensions in future work; the harmonic approximation adopted in the present work will probably lead to a systematic underprediction of crossover temperatures, since the softer modes of higher energy minima are expected to be more anharmonic. Corresponding shifts in heat capacity peaks and melting points have been analysed in previous work [17,27].

As each permutation-inversion isomer of a given structure, $Z_i(T)$ is identical, the sum above can be written as the product of the partition function for one of these isomers and a factor n_i, which accounts for the permutation-inversion degeneracy. For a system containing N_A atoms of element A, N_B atoms of element B, etc., $n_i = 2N_A! \, N_B! \, N_C! \ldots /o_i$, where o_i is the order of the point group of minimum i [18,31–33]. Ignoring the rotational contribution to the partition function, which does not usually vary significantly between different isomers of the same atomic cluster, and making a harmonic approximation to the vibrational component of the partition function, one can write

$$Z(T) = \sum_i \frac{n_i \exp[-\beta V_i]}{(\beta h \bar{\nu}_i)^\kappa}, \qquad (2)$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, κ is the number of non-zero eigenvalues for the Hessian matrix, $\bar{\nu}_i = (\prod_{j=1}^{\kappa} \nu_i^j)^{1/\kappa}$ is the geometric mean vibrational frequency of minimum i, with ν_i^j the normal mode frequency of the j-th mode in minimum i, and V_i is the potential energy of the minimum i. The approximate free energy of minimum i is then given by $F_i(T) = -k_B T \ln Z_i(T)$. It is apparent from Eq. (2) that the effect of temperature will come into play most rapidly for competing structures with different point group symmetries or mean vibrational frequencies. The rotational partition function can easily be included if required, assuming a rigid rotor model for the density of states, but we have checked that it is not important for the examples considered below.

In the free energy basin-hopping (FEBH) framework introduced here the local free energies of the two minima are used in the accept/reject test for step-taking between local minima. The key formula is

$$F_{\text{new}}(T) - F_{\text{old}}(T) = V_{\text{new}} - V_{\text{old}} + k_B T \ln \frac{o_{\text{new}} \bar{\nu}_{\text{new}}^\kappa}{o_{\text{old}} \bar{\nu}_{\text{old}}^\kappa}. \qquad (3)$$

The product of vibrational normal mode frequencies can be obtained from the product of non-zero eigenvalues of the mass-weighted Hessian.

The FEBH approach has now been implemented in the GMIN programme [34]. To illustrate the utility of this scheme and provide proof of principle we have tested it for various atomic and colloidal clusters, along with some peptides to provide examples for biomolecules. In future work we will report on applications to much larger biomolecules, with particular focus on systems involving ligand binding. In particular, we will optimise the efficiency of the local free energy calculations, which involve the additional expense entailed by a normal mode analysis. Although it is also possible to calculate the local harmonic free energies for a sample of potential energy minima obtained after a conventional basin-hopping run, we find examples where FEBH locates the local free energy global minimum using fewer steps, on average.

2. Applications

Lennard-Jones (LJ) clusters have been extensively studied for benchmarking global optimisation methods [15,35]. The Lennard-Jones potential can be written as $V(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, where ϵ is the pair well depth and σ is the separation r for which $V(r)$ vanishes. Figure 1a shows the free energy as a function of temperature for the minima of the seven-atom Lennard-Jones cluster, LJ_7, for which there are only four distinct structures. The global potential energy minimum corresponds to a pentagonal bipyramidal structure (D_{5h}), and the other three local minima are the capped octahedron (C_{3v}), tricapped tetrahedron (C_{3v}), and skew structure (C_2) in ascending order of potential energy [36,37]. The respective orders of the point group are 20, 6, 6, and 2. It is evident that the skew structure, which possesses the lowest symmetry, becomes relatively more stable as temperature increases, and is the second most stable structure after the pentagonal bipyramid at $T = 0.25$ in units of ϵ/k_B.

It is instructive to compare the case of a seven-particle colloidal cluster, bound by a short-ranged Morse potential, M_7. The Morse potential is given by $V(r) = \epsilon \exp[\rho(1 - r/r_e)](\exp[\rho(1 - r/r_e)] - 2)$, where ϵ is the pair well depth, r_e is the pair equilibrium separation, and ρ is a parameter that controls the range of the potential. With the range parameter $\rho = 30$, the Morse potential was found to provide a quantitative description of occupation probability distribution for competing structures of polystyrene microspheres as measured by optical microscopy [38,39]. For $\rho = 30$, the global minimum on the PES is also the pentagonal bipyramid, as shown in Figure 2a, whilst three other minima have the same number of nearest neighbour contacts and are of comparable potential energies. Our global optimisation runs confirm that the M_7 cluster in this case has more minima as compared to the LJ_7 cluster, which is due to the short-ranged pair interactions for $\rho = 30$. The relative stabilities of these structures are therefore more sensitive to
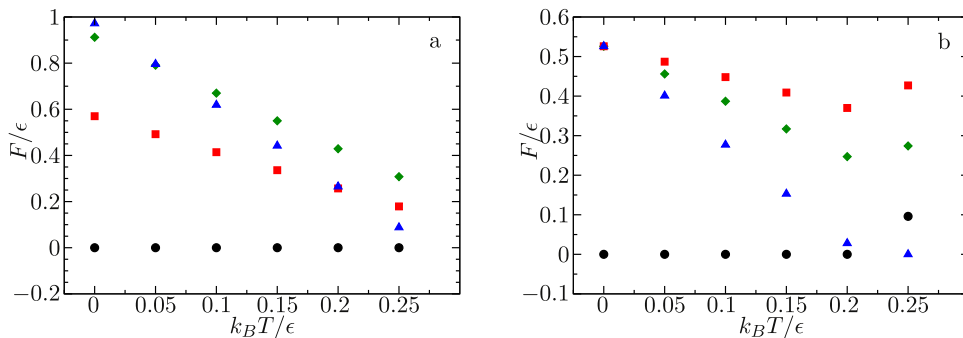
Figure 1. Free energy of the minima as a function of temperature for the Lennard-Jones cluster LJ_7 and the Morse cluster M_7 with range parameter $\rho = 30$. (a) The LJ_7 cluster has four distinct minima: the pentagonal bipyramid (circle), capped octahedron (square), tricapped tetrahedron (diamond), and skewed structure (triangle). (b) The M_7 cluster has four low-lying minima with the same number of nearest neighbour contacts: pentagonal bipyramid (circle), capped octahedron (square), tricapped tetrahedron (diamond), and bicapped trigonal bipyramid (triangle). The free energies are presented relative to that of the most stable structure, which is set to zero at all temperatures.