MOLECULAR PHYSICS, 2002, VOL. 100, NO. 2, 3285–3305

Taylor & Francis
Taylor & Francis Group

Discrete path sampling

DAVID J. WALES
University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, UK

(*Received 13 April 2002; accepted 20 June 2002*)

A theoretical framework is developed for the calculation of rate constants by sampling connected pathways composed of local minima and transition states that link them together. The theory is applicable to two-state or effective two-state systems and is applied to permutational or morphological isomerization in a two-dimensional cluster of seven Lennard-Jones atoms, water clusters containing eight and nine molecules, and a cluster of 38 Lennard-Jones atoms, which exhibits a double funnel energy landscape.

1. Introduction

Many dynamical processes of great physical importance are too slow to study by conventional simulation techniques with current computer technology. Hence there is much interest in techniques to calculate the rate constants of such 'rare events' more efficiently. Significant progress was made by separating the problem into the calculation of a free energy barrier and a dynamical factor corresponding to the crossing rate [1, 2]. The probability of finding the system at the top of the free energy barrier is an equilibrium quantity, and can be calculated by either Monte Carlo or molecular dynamics techniques. To obtain good statistics umbrella sampling is usually employed [3]. To calculate the crossing rate constrained molecular dynamics simulations may be used, such as the 'blue moon ensemble' [4, 5]. In a recent case study the above techniques proved sufficiently powerful to calculate the homogeneous nucleation rate in a moderately supercooled Lennard-Jones system [6]. However, there are also problems that must be overcome in such applications: successful sampling may require the construction of a suitable reaction coordinate or biasing potential by trial and error, and the biasing potential may be very steep for processes with large free energy barriers. One difficulty that has now been overcome is the question of how to sample dynamical pathways between two stable states in a statistically correct fashion, without prior knowledge of a free energy transition state [7–20]. Other recent work includes algorithms for enhanced sampling of a given reaction coordinate [21], as well as kinetic Monte Carlo (KMC) [22–25] and various other techniques that utilise information about stationary points on the potential energy surface (PES) [26–30].

In the present contribution we present the discrete analogue of dynamical path sampling, where pathways correspond to sequences of local minima and transition states, and thermodynamics and dynamics are calculated using model densities of states. Here we define a transition state as a point where the gradient vanishes and the Hessian matrix of second derivatives has exactly one negative eigenvalue [31]; a local minimum has no negative Hessian eigenvalues. Connections between the local minima are defined by calculating steepest-descent pathways for each transition state, or approximations to them.

Both thermodynamic and dynamic properties can be calculated using databases of local minima and transition states [32]. The problem with this approach is that the number of local minima on the PES grows exponentially with system size [33–35], and so sampling rapidly becomes an issue for complex systems. Previous schemes to obtain statistical samples of local minima and transition states have generally used thermodynamic criteria to decide which parts of the landscape have the highest weights. Obviously one can only sample a small fraction of the local minima in a large system, just as only a small fraction of phase space is sampled in Monte Carlo or molecular dynamics simulations. Samples of local minima can be corrected for incompleteness using a reweighting scheme that is related to the histogram Monte Carlo approach [36–38]. Instead of determining the configurational density of states from the canonical potential energy distribution, a density of minima is obtained from the occupation probability of the different local minima [32, 39]. Combined with the superposition approximation [32], where the partition function is written as a weighted sum over the contributions of local minima, this approach provides a powerful way to calculate approximate equilibrium thermodynamic properties in quasi-ergodic systems [40]. If sophisticated techniques, such as parallel [41] or entropic [42] tempering, are unable to achieve equilibrium sampling,

Molecular Physics ISSN 0026–8976 print/ISSN 1362–3028 online
http://www.tandf.co.uk/journals
DOI: 10.1080/00268970210162691

then the superposition approach may be the only way to proceed.

Dynamical properties, based on large samples of minima and transition states and master equation dynamics, have been reported for a variety of systems [43–62]. However, these calculations may underestimate the rates if the samples omit higher energy local minima and transition states that are critical for the dynamics of interest, but contribute little to the equilibrium thermodynamic properties. This problem can be overcome using the discrete path sampling approach described in § 2. Visualisation of potential energy surfaces using disconnectivity graphs [50] has helped to provide detailed insight into thermodynamics and dynamics at a microscopic level [32]. The discrete path sampling approach also provides a means to reduce the information encoded in such graphs to a manageable level for large systems, and will select a representation that is most appropriate for considering dynamics.

2. Theory

We start from the master equation view of dynamics in terms of the time evolution of the occupation probability for finding the system in local minimum α at time t:

$$\frac{\mathrm{d}P_\alpha(t)}{\mathrm{d}t} = \sum_{\beta\neq\alpha}[k_{\alpha\beta}P_\beta(t) - k_{\beta\alpha}P_\alpha(t)], \qquad (1)$$

where $k_{\alpha\beta}$ is the first order rate constant for transitions from minimum β to minimum α; the restriction $\beta \neq \alpha$ removes processes that do not affect $P_\alpha(t)$, such as degenerate rearrangements [63] between permutational isomers of the same structure. This linear master equation is based on the assumption that the system has time to equilibrate sufficiently between transitions to lose its memory of previous jumps, i.e. the dynamics must be Markovian. We expect this approximation to become increasingly poor as the temperature increases; this is the limit where standard simulations should be feasible.

An analytical solution can be written for equation (1) as a sum of decaying exponentials [64, 65]. It only makes sense to think in terms of phenomenological rate constants for interconversion of two states if all the exponentials bar one decay rapidly on the experimental time scale. This condition is equivalent to the alternative view that the reaction time must be much longer than the molecular relaxation time in order for the reactive flux to exhibit a plateau region that defines the rate constant [2, 10]. Although path sampling does not require prior knowledge of a reaction coordinate, we do need to be able to distinguish the two states between which we wish to calculate the rate constant. Let us call these states A and B, and assume that we have an order parameter that enables us to assign local minima to one or the other. Single exponential decay from a non-equilibrium distribution will only be observed if there is local equilibrium among the minima belonging to state A, $\{a\}$, and among the minima belonging to state B, $\{b\}$. With

$$P_A(t) = \sum_{a\in A} p_a(t) \qquad \text{and} \qquad P_B(t) = \sum_{b\in B} p_b(t), \qquad (2)$$

the occupation probabilities of individual minima corresponding to the local equilibria in states A and B are defined by

$$p_a(t) = \frac{p_a^{\mathrm{eq}} P_A(t)}{P_A^{\mathrm{eq}}} \qquad \text{and} \qquad p_b(t) = \frac{p_b^{\mathrm{eq}} P_B(t)}{P_B^{\mathrm{eq}}}, \qquad (3)$$

where the 'eq' superscript stands for 'equilibrium'. Equation (1) is then easily manipulated to give

$$\frac{\mathrm{d}P_A(t)}{\mathrm{d}t} = k_{AB}P_B(t) - k_{BA}P_A(t)$$

and

$$\frac{\mathrm{d}P_B(t)}{\mathrm{d}t} = k_{BA}P_A(t) - k_{AB}P_B(t), \qquad (4)$$

where

$$k_{AB} = \frac{1}{P_B^{\mathrm{eq}}} \sum_{a\in A}\sum_{b\in B} k_{ab}\, p_b^{\mathrm{eq}}$$

and

$$k_{BA} = \frac{1}{P_A^{\mathrm{eq}}} \sum_{a\in A}\sum_{b\in B} k_{ba}\, p_a^{\mathrm{eq}}. \qquad (5)$$

Under these conditions equation (5) shows that the phenomenological rate constants can be written as a sum over all the rate constants for elementary well-to-well transitions $a \quad b$ and $b \quad a$ that lie on the boundary of the two regions. Each elementary rate constant is weighted by the conditional equilibrium probability of finding the system in the starting minimum, given that we are in the appropriate A or B state.

We can take the above analysis one stage further by allowing for minima that cannot be classified as either A or B in character. This extension may be useful for several reasons. Firstly, in the case of a high free energy barrier it may not be reasonable to assume that all $a \in A$ and all $b \in B$ are in local equilibrium. Secondly, it may be unrealistic to classify all minima as either A or B in character, and we then need a theory where the rate constants do not depend sensitively upon the position of the A/B boundary. We therefore admit intermediate minima, i, belonging to a third set, I. Effective two-state $A \leftrightarrow B$ dynamics is recovered if we apply the steady-state approximation to the occupation probability of each intermediate minimum i, assuming that these minima only have low occupation probabilities. Then we can write:

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_\alpha k_{i\alpha}\, p_\alpha(t) - p_i(t) \sum_\beta k_{\beta i} \approx 0,$$

so that
$$p_i(t) = \frac{\sum_\alpha k_{i\alpha}\, p_\alpha(t)}{\sum_\beta k_{\beta i}}. \tag{6}$$

Using equation (6) we can write an expression for $\mathrm{d}P_A(t)/\mathrm{d}t$ and substitute for the probability of any intermediate minimum that appears. Each substitution introduces a linear combination of probabilities $p_i(t)$, $p_a(t)$ and $p_b(t)$. Hence $\mathrm{d}P_A(t)/\mathrm{d}t$ and $\mathrm{d}P_B(t)/\mathrm{d}t$ can be written as sums over all possible paths that start and finish on the boundaries of the A and B regions:

$$\begin{aligned}
\frac{\mathrm{d}P_A(t)}{\mathrm{d}t} = & \sum_a \sum_{a'} \frac{k_{ai_1} k_{i_1 i_2} \ldots k_{i_n a'}\, p_{a'}(t)}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \\
& + \sum_a \sum_b \frac{k_{ai_1} k_{i_1 i_2} \ldots k_{i_n b}\, p_b(t)}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \\
& - \sum_a \sum_i k_{ia}\, p_a(t), \\
\frac{\mathrm{d}P_B(t)}{\mathrm{d}t} = & \sum_b \sum_{b'} \frac{k_{bi_1} k_{i_1 i_2} \ldots k_{i_n b'}\, p_{b'}(t)}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \\
& + \sum_b \sum_a \frac{k_{bi_1} k_{i_1 i_2} \ldots k_{i_n a}\, p_a(t)}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \\
& - \sum_b \sum_i k_{ib}\, p_b(t),
\end{aligned} \tag{7}$$

where $a \quad b$ indicates that the sum is over all possible paths that begin from a minimum $b \in B$ and end in a minimum $a \in A$, passing only through minima $i \in I$. Since $\mathrm{d}P_I(t)/\mathrm{d}t = 0$, because of the steady state assumption for all the intermediate minima, we still have $\mathrm{d}P_A(t)/\mathrm{d}t = -\mathrm{d}P_B(t)/\mathrm{d}t$, and so we deduce that equation (4) still holds with

$$\begin{aligned}
k_{AB} &= \frac{1}{P_B^{\mathrm{eq}}} \sum_a \sum_b \frac{k_{ai_1} k_{i_1 i_2} \ldots k_{i_n b}\, p_b^{\mathrm{eq}}}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}}, \\
k_{BA} &= \frac{1}{P_A^{\mathrm{eq}}} \sum_b \sum_a \frac{k_{bi_1} k_{i_1 i_2} \ldots k_{i_n a}\, p_a^{\mathrm{eq}}}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}},
\end{aligned} \tag{8}$$

where we have again applied the local equilibrium conditions, e.g. $p_a(t) = p_a^{\mathrm{eq}} P_A(t)/P_A^{\mathrm{eq}}$. Note that the first and third terms in the expression for $\mathrm{d}P_A(t)/\mathrm{d}t$ in equation (7) must combine to leave only $b \quad a$ paths by comparison with the expression for $\mathrm{d}P_B(t)/\mathrm{d}t$, and vice versa. $b \quad a$ paths arise from putting terms in $p_a(t)$ for $\mathrm{d}P_A(t)/\mathrm{d}t$ over a common denominator, because the sums in the denominator of the $a \quad a'$ term include rate constants from b to i minima. Direct A to B connections are also included in the more general expressions in (8).

When an intermediate region is included the sum over rate constants for processes crossing the A/B boundary in equation (5) is replaced by a sum over paths starting and finishing on the A/I and B/I boundaries. We will refer to the corresponding minimum−transition state−minimum ...−minimum sequences as discrete paths. The sums over rate constants for individual local minimum to local minimum transitions are weighted by the conditional occupation probability for the starting minimum given that the system is in the appropriate A or B region. These rate constant expressions, although discretised in terms of individual transitions between minima, will still contain an unfeasibly large number of contributions for large systems. We must therefore determine how to sample them effectively.

Since detailed balance applies, so that $k_{AB}/k_{BA} = P_A^{\mathrm{eq}}/P_B^{\mathrm{eq}}$, it is only necessary to calculate one of the rate constants along with the equilibrium occupation probabilities. These probabilities can be obtained from free energies or entropies for the canonical and microcanonical ensembles, respectively. In fact, detailed balance is obeyed for each specific $a - b$ path because

$$\begin{aligned}
& \frac{1}{P_B^{\mathrm{eq}}} \frac{k_{ai_1} k_{i_1 i_2} \ldots k_{i_n b}\, p_b^{\mathrm{eq}}}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \\
& \qquad = \frac{P_A^{\mathrm{eq}}}{P_B^{\mathrm{eq}}} \left(\frac{1}{P_A^{\mathrm{eq}}} \frac{p_a^{\mathrm{eq}} k_{i_1 a} k_{i_2 i_1} \ldots k_{bi_n}}{\sum_{\alpha_1} k_{\alpha_1 i_1} \sum_{\alpha_2} k_{\alpha_2 i_2} \ldots \sum_{\alpha_n} k_{\alpha_n i_n}} \right),
\end{aligned}$$

or $k_{a \quad b} P_B^{\mathrm{eq}} = k_{b \quad a} P_A^{\mathrm{eq}}$ for any path. This result provided a useful consistency check in the numerical calculations described later, and also shows that it is only necessary to sample the individual paths in one sense for cases where all the local minima in the A and B sets are known. Examples where one or both of these sets contain an exponentially large number of minima will be considered elsewhere.

Calculating a particular $a - b$ path entails a number of local minimisations and transition state searches, and may be more computationally expensive than running a single molecular dynamics trajectory. However, the discrete path subsumes all the dynamical trajectories that pass through the catchment basins of the corresponding sequence of local minima, where a catchment basin is a well-defined region of configuration space from which steepest-descent paths lead to one particular minimum

[66, 67]. A discrete $a-b$ path is also not subject to trapping; transition states involving very small rate constants will make a negligible contribution to k_{AB} and k_{BA}, but can be calculated just as easily as those that make large contributions.

2.1. *Recrossings*

Let us define the number of steps in a given discrete $a-b$ path as the number of transition states in the sequence, which is one less than the total number of local minima. A path with three or more steps can be augmented by recrossings of any transition state involving only intermediate minima, to produce an infinite family of discrete paths that link the same a/b pair via the same set of intermediate stationary points. The contribution to k_{AB} and k_{BA} diminishes with each recrossing: for one recrossing of transition state i, linking intermediate minima i and $i+1$, the corresponding contribution to the sum over paths gains an extra factor of

$$\frac{k_{i+1\,i}\,k_{i\,i+1}}{\sum_{\alpha} k_{\alpha\,i+1}\sum_{\beta} k_{\beta\,i}} < 1, \tag{9}$$

because at least one other connection must be known for each of the two minima if they are part of an $a-b$ path, and the sums are over all such connections and also contain $k_{i\,i+1}$ and $k_{i+1\,i}$, respectively. For short paths it is not difficult to sum all the recrossing contributions analytically. However, the resulting formulae rapidly become complicated because of the combinatorial factors that must be included when more than one transition state can be recrossed and the order in which the recrossings occur must be accounted for. Fortunately it is still possible to sum all the recrossing contributions so that only $a-b$ paths without recrossings need be sampled. The computational cost involved is negligible compared to the time spent finding the pathways. The extra contribution of recrossings to k_{ab} and k_{ba} paths varies from about 10% to a factor of ten in the examples considered in the present work.

We can calculate the contribution of all possible $a-b$ paths for a given sequence of minima and transition states using results from graph theory. Suppose we have a set of n objects and knowledge of how they are connected together; the objects and connections are termed *vertices* and *edges* of a graph [68]. We can define an adjacency matrix, $\mathbf{A}$, whose elements are

$$A_{\alpha\beta} = \begin{cases} 1, & \alpha \text{ and } \beta \text{ are connected by an edge,} \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

It is not difficult to prove that the number of paths of length k steps between vertices α and β is $[\mathbf{A}^k]_{\alpha\beta}$ [68]. To solve the present problem for a path where each intermediate minimum occurs only once we construct the weighted adjacency matrix with elements

$$A_{i_\alpha i_\beta} = \begin{cases} k_{i_\alpha i_\beta}/\sum_\gamma k_{\gamma i_\alpha}, & |i_\alpha - i_\beta| = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

so that

$$\mathbf{A} = \begin{Bmatrix} 0 & k_{i_1 i_2}/\sum_\gamma k_{\gamma i_1} & 0 & 0 & 0 & \dots \\ k_{i_2 i_1}/\sum_\gamma k_{\gamma i_2} & 0 & k_{i_2 i_3}/\sum_\gamma k_{\gamma i_2} & 0 & 0 & \dots \\ 0 & k_{i_3 i_2}/\sum_\gamma k_{\gamma i_3} & 0 & k_{i_3 i_4}/\sum_\gamma k_{\gamma i_3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{Bmatrix} \tag{12}$$

The dimension of $\mathbf{A}$ is equal to the number of intermediate minima, because only recrossings of transition states that link intermediate minima are allowed. The minimum dimension is two, for there can be no recrossings with less than two intermediate minima. The contributions of a given $a-b$ path to k_{AB} and k_{BA} are then

$$k_{ab} = \frac{p_b^{\mathrm{eq}}}{P_B^{\mathrm{eq}}} k_{a i_1} \frac{k_{i_n b}}{\sum_\gamma k_{\gamma i_n}} \sum_{p=n-1}^{\infty} [\mathbf{A}^p]_{1\,n},$$

$$k_{ba} = \frac{p_a^{\mathrm{eq}}}{P_A^{\mathrm{eq}}} k_{b i_n} \frac{k_{i_1 a}}{\sum_\gamma k_{\gamma i_1}} \sum_{p=n-1}^{\infty} [\mathbf{A}^p]_{n\,1}, \tag{13}$$

where the sums are over p odd or even for n even or odd, respectively, and n is the number of intermediate minima for the given $a-b$ path, which is equal to the number of transition states minus one. The ratios of probabilities are, of course, equal to the ratios of the corresponding partition functions. In practice the sums over p typically converge for $p < 100$; for a few longer pathways we found that numerical precision was lost in evaluating the sums, but the contribution of such pathways to the quantities of interest was always negligible.

It is possible to rewrite the expressions in (13) in terms of the eigenvalues of $\mathbf{A}$. Although $\mathbf{A}$ is an asymmetric matrix, the tridiagonal form in equation (12) has real eigenvalues. This result follows because $\mathbf{A}$ is similar to a symmetric matrix $\mathbf{DAD}^{-1}$, with $\mathbf{D}$ a diagonal matrix, so long as [69]

$$A_{i\,i+1} A_{i+1\,i} > 0, \tag{14}$$

which is always true in the present case.

If the left and right eigenvectors of $\mathbf{A}$ are collected in the columns of matrices $\mathbf{L}$ and $\mathbf{R}$, respectively, then we can define the eigenvector normalisation so that

$$\left[\mathbf{L}^T \mathbf{A} \mathbf{R}\right]_{\alpha\beta} = \lambda_\alpha \delta_{\alpha\beta}, \qquad \text{and} \qquad \left[\mathbf{L}^T \mathbf{A}^p \mathbf{R}\right]_{\alpha\beta} = \lambda_\alpha^p \delta_{\alpha\beta}, \tag{15}$$

where superscript T denotes the transpose and λ_α is eigenvalue α. Writing $\mathbf{A}^p = \mathbf{R}\Lambda^p\mathbf{L}^T$, where $\Lambda_{\alpha\beta} = \lambda_\alpha \delta_{\alpha\beta}$, gives

$$\sum_{p=n-1}^{\infty} [\mathbf{A}^p]_{\alpha\beta} = \sum_{q=1}^{n} R_{\alpha q} L_{\beta q} \frac{\lambda_q^{n-1}}{1-\lambda_q^2}, \quad \text{if} \quad |\lambda_\alpha| < 1, \ \forall\ \alpha. \tag{16}$$

Although the resulting expressions for k_{AB} and k_{BA} are neater in this form, problems of numerical precision also arose for a few long paths that make a negligible contribution. In the present work there was little to choose between the two alternative routes to the recrossing calculation.

It is important to note that the recrossing factors discussed above are different from the transmission coefficients that arise in conventional dynamical calculations [2]. In the present framework recrossings that reduce the usual transition state theory rate constant could be included in the individual $k_{\alpha\beta}$ corresponding to transitions between local minima [2]. The discrete path sampling approach presented above is independent of the methods used to calculate the probabilities p_α^{eq} and the individual $k_{\alpha\beta}$. In the present work we have used simple approximations to the latter quantities, based upon harmonic densities of states and conventional transition state theory for the rate constants. However, the discrete path sampling approach is by no means limited to such a framework, and anharmonic classical or quantum densities of states could be used [70–72], along with more sophisticated calculations of the individual $k_{\alpha\beta}$, up to and including approaches based on molecular dynamics.

3. Partition functions and rate constants

In this first application of discrete path sampling we employ the simplest harmonic approximation for the local densities of states along with the corresponding formulation of transition state theory. The microcanonical and canonical results are:

$$\Omega_\alpha(E) = \frac{n_\alpha (E-E_\alpha)^{\kappa-1}}{\Gamma(\kappa)(h\overline{\nu}_\alpha)^\kappa}$$

and

$$Z_\alpha(T) = \frac{n_\alpha \exp(-\beta E_\alpha)}{(\beta h\overline{\nu}_\alpha)^\kappa}, \tag{17}$$

$$k_{\alpha\gamma}(E) = \frac{o_\gamma}{o_\dagger} \frac{\bar{\nu}_\gamma^\kappa}{\bar{\nu}_\dagger^{\kappa-1}} \left(\frac{E-E_\dagger}{E-E_\gamma}\right)^{\kappa-1},$$

and

$$k_{\alpha\gamma}(T) = \frac{o_\gamma}{o_\dagger} \frac{\bar{\nu}_\gamma^\kappa}{\bar{\nu}_\dagger^{\kappa-1}} \mathrm{e}^{-(E_\dagger - E_\gamma)\beta},$$

where E_α is the potential energy of minimum α, $\overline{\nu}_\alpha$ is the geometric mean vibrational frequency, κ is the number of vibrational degrees of freedom, $\beta = 1/kT$ and Γ is the gamma function. n_α is the number of distinct (non-superimposable) permutation-inversion isomers of minimum α; for an atomic cluster containing N atoms of the same element $n_\alpha = 2N!/o_\alpha$, where o_α is the order of the point group of α. For a cluster in field-free space, $\kappa = 3N-6$; we have ignored the rotational and translational degrees of freedom since we generally consider systems with zero linear and angular momentum. The † subscript denotes the corresponding quantities for the transition state in question.

4. Methods

Most of the algorithms used to calculate stationary points and pathways in the present work have been described before [32], so only a brief overview will be given here. All energy minimisations were performed using a modified version of the method described by Liu and Nocedal [73]; we found it more efficient to remove the line searches from the original procedure and include a dynamical scaling scheme for the step length. Energy minimisation was also used instead of calculating true steepest-descent pathways, with an initial step of 0.04 units along the Hessian eigenvector corresponding to the unique negative eigenvalue at each transition state. We have previously checked that the resulting connections are almost always the same. Full details of the convergence criteria and other parameters involved in the calculations are given in the Appendix.

As in previous work we employed hybrid eigenvector-following techniques to seek a transition state from a single starting point [32, 74, 75]. We will refer to such searches as 'single-ended'. When intermediate local minima are involved the expressions for k_{AB} and k_{BA} contain sums of rate constants for the flux out of these minima. The same sum arises when calculating the waiting time in kinetic Monte Carlo simulations [22–25]. Intermediate minima on an $a-b$ path must already have at least two separate connections to adjacent minima. For each new minimum found we conducted single-ended transition state searches to locate a minimal number of extra connections; the hard sphere collision model used in previous work [76] was employed to generate initial guesses for these transition state searches. In practice we found that the rate constant sums were dominated by a relatively small number of contributions, which were usually found early in the searches. One attractive feature of discrete path sampling is that the contribution of each $a-b$ path to k_{BA} and k_{BA} can be recalculated using all the connections collected for the intermediate local minima during the path sampling run. Additional connections will always reduce the rates, but the effect of such corrections on the final results was always rather small for the examples considered in the present work.

It was also necessary to locate pathways between particular local minima, and we will refer to these as 'double-ended' searches. As for single-ended searches a

variety of algorithms have been proposed for this purpose. We used energy minimisation to calculate pathways, as described above, and single-ended hybrid eigenvector-following [32, 74, 75] to locate transition states. The initial guess for each of these transition state searches was provided by the modified 'nudged elastic band' (NEB) method [25, 77–79]. Although complete paths involving more than one transition state can be characterised with this method, we found that the higher energy regions of the path converged much faster than the low energy regions in our implementation. We therefore employed NEB calculations with a small number of images and a limited number of iterations, and then used the highest energy image as the starting point for a hybrid eigenvector-following transition state search. A flow-diagram for the overall double-ended procedure is given in figure 8 in the Appendix, along with all the parameters used.

The relation between single- and double-ended algorithms for finding transition states on single pathways is somewhat analogous to the relation between the kinetic Monte Carlo (KMC) approach and the discrete path sampling method described in the present work. In KMC simulations information from the potential energy surface about connections from the current local minimum is used to advance the system in time. In contrast, the discrete path sampling approach can be used when the final and initial states of the system are specified in advance, and is therefore a 'double-ended' philosophy.

The remainder of the new algorithm is concerned with the business of sampling. The precise details should probably be adapted to the system in question. In all the cases considered in the present work the A and B regions could be specified from the outset in terms of finite sets of particular local minima. Given just one path between A and B, calculated by a single double-ended search between one minimum selected from A, and one from B, how do we generate new paths? The analogous problem has already been addressed in dynamical path sampling studies [8–18, 20] and some of the results are likely to be useful in discrete path sampling. For example, looking at the formula for k_{AB} in equation (8) we note that an importance sampling scheme can be devised using p_b^{eq} (or p_a^{eq} for k_{BA}).

Although the formulae for k_{AB} and k_{BA} involve formally infinite sums over all possible paths, we found that only a few paths made significant contributions for the examples considered in the present work. Hence it is most important to identify these fastest paths in the fewest possible steps. In this respect the evaluation of k_{AB} and k_{BA} is more like an exercise in global optimisation, although we will generally require more than just the fastest path.

First we need a way to produce a new path from a given initial path. There are many ways that this might be done, and after a few trials we settled on the following procedure. First, a local minimum from the current path (including the a and b endpoints) was chosen at random. This minimum was replaced by a different one, either from the database of known connections, or by perturbing the geometry and minimising until a new structure was found. Double-ended searches were then conducted to link this replacement minimum to others in the original path that lay a given number of steps away. Once the connections were finished we calculated k_{ab} and k_{ba} if a new path had been found. The present results were obtained by performing a fixed maximum number of perturbations, n_{pert}, for the fastest n_{max} paths in the database. These fastest paths change throughout the run—if a faster path was found following a perturbation then it became the current path. Hence, at any stage we were always perturbing the fastest path for which fewer than n_{pert} perturbations had been considered. A flow-diagram of this procedure along with details of the various parameters used is given in figure 9 in the Appendix.

All the $a-b$ paths found were used in the final evaluation of k_{AB} and k_{BA}, together with the updated sums of rate constants out of each intermediate minimum. Once the paths are known they may also be used to calculate k_{AB} and k_{BA} as a function of temperature (canonical ensemble) or total energy (microcanonical ensemble). This procedure requires only a few seconds of computer time, but carries with it the assumption that the same paths dominate over the given temperature range. If new runs are instead conducted at different temperatures, then all the saved local minima and transition states from a previous run can be used to speed up the calculation. It is easy to imagine discrete path sampling schemes for parallel computers that would exploit the above algorithms efficiently. However, all the present results were obtained rapidly enough using only a single processor.

5. Results for two-dimensional LJ_7

Contributions to the permutational isomerization rate for the global minimum of the seven-atom, two-dimensional Lennard-Jones cluster, LJ_7, were calculated by Dellago, Bolhuis and Chandler using a dynamical path sampling approach [9]. This system has since served as a test case for an alternative approach to rare event dynamics [19]. To treat permutational isomerization within the discrete path sampling approach the atom initially in the middle of the centred hexagon global minimum was tagged by assigning it a different mass. This mass enables permutational isomers where the tagged atom occupies inequivalent positions to be

Table 1. Local minima and transition states for two-dimensional LJ_7. With a tagged atom in the central position the global minimum has a six-fold rotation axis; this is the only relevant symmetry if the atoms are constrained to a plane.

Minimum	Energy	Transition state	Energy
1	−12.5348	1	−11.0402
2	−11.5013	2	−11.0373
3	−11.4769	3	−10.9352
4	−11.4034	4	−10.9198
		5	−10.8985
		6	−10.8825
		7	−10.8413
		8	−10.8073
		9	−10.7987
		10	−10.7980
		11	−10.7714
		12	−10.5515
		13	−10.5469
		14	−10.4827
		15	−10.4547
		16	−10.4128
		17	−10.4031
		18	−10.3881
		19	−10.3760

identified from the principal moments of inertia. Tagging one atom also assists in the assignment of appropriate point group symmetries, which are pure rotation groups in two-dimensional space, and hence ensures that the correct reaction path degeneracies were used. For example, the forward and backward rates must differ by a factor of six, because there are six times as many distinct permutational isomers of the global minimum with a tagged particle on the outside than when the tagged particle is in the middle of the cluster. The normal mode frequencies must, however, be calculated using unit mass for each atom. Here and in § 8 we will employ reduced units for the LJ potential, where ϵ is the unit of energy, σ is the unit of distance, and $\sqrt{\epsilon/m\sigma^2}$ is the unit of frequency, where m is the atomic mass.

The potential energy surface for this system supports four local minima and nineteen true transition states (table 1), excluding permutation-inversion isomers. Discrete path sampling runs were started from an initial five-step path, which was calculated according to the prescription in the Appendix. A total of 636 distinct paths were found, and the fastest four are illustrated in figure 1. Each one is a three-step path, and the canonical contributions to the rate constant for $kT/\epsilon = 0.05$ are 1.3×10^{-14}, 8.1×10^{-15}, 7.5×10^{-15} and $6.2 \times 10^{-15}\sqrt{\epsilon/m\sigma^2}$ for the direction in which the tagged atom moves from the centre of the cluster to the outside. The next-largest contribution is about a factor of two smaller, and the fastest four paths contribute about 58% to the total rate constant of 6.0×10^{-14}, indicating that there are a number of competitive alternatives at this temperature. The fastest ten paths make up about 74% of the total rate constant at this temperature. All the reverse rates are a factor of six slower for any canonical temperature or microcanonical total energy, as required by symmetry. Results for the rates as a function of temperature or total energy are omitted for brevity.

Three of the four fastest paths were illustrated in the previous dynamical path sampling study—the other one simply involves two of the same rearrangements in a different order. They all involve three successive diamond-square-diamond processes [80], as expected for the two-dimensional analogue of a deltahedral cluster [81]. The relatively large number of processes that contribute reflects the lack of one particularly favourable mechanism for this isomerization.

We also performed discrete path sampling calculations corresponding to the simpler rearrangements for which rate constants were reported in previous work (table 2). All the present results are within an order of magnitude of the more accurate dynamical path sampling calculations [9].

6. Results for $(H_2O)_8$

For this system, and for the LJ_{38} cluster considered in § 8, databases of local minima and transition states were already available from previous work [92, 99]. We first used these databases in rate constant calculations without any further searches for pathways. The databases were analysed to find pathways between the specified A and B states for increasing numbers of steps, and the contributions to k_{AB} and k_{BA} can be compared with previous results from dynamical path sampling or master equation dynamics.

For $(H_2O)_8$ we consider isomerization between the two lowest-energy isomers, which are cuboids with D_{2d} and S_4 point group symmetry. This system provides an interesting example of competition between two low-energy structures, and has been studied in a number of previous investigations [16, 82–87]. Experimentally, the two isomers have been identified both as free clusters [88] and attached to a benzene molecule [89]; the octamer has also been found in a crystal structure [90].

For comparison with a previous dynamical path sampling study [16] we used the microcanonical ensemble and the TIP4P rigid body intermolecular potential [91]. The number of intermolecular vibrational degrees of freedom is therefore $\kappa = 6 \times 8 - 6 = 42$. The rate constant for interconversion of the S_4 to the D_{2d} isomer at a total energy of −253.13 kJ/mol was

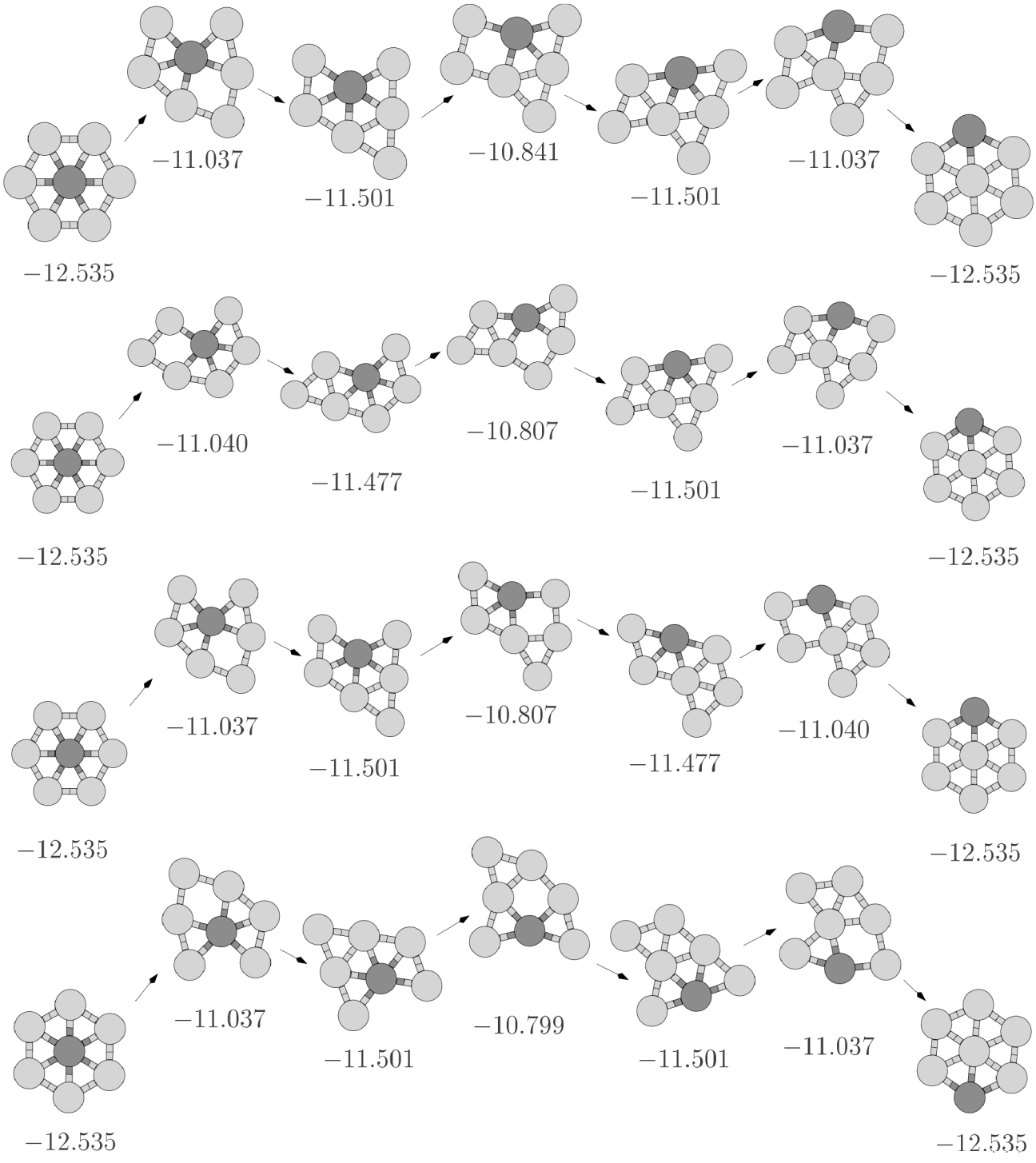

Figure 1. The four fastest paths at $kT/\epsilon = 0.05$ for permutational isomerization of the global minimum of two-dimensional LJ_7. The tagged atom is highlighted.

calculated as $5 \times 10^4\,s^{-1}$ in the previous dynamical path sampling study [16]. This result appears to correspond to the interconversion rates for two specific permutational isomers, and so for comparison we set all the point group orders in (17) to unity. The total $k_{D_{2d}\ S_4}$ and $k_{S_4\ D_{2d}}$ rates are four and eight times greater, respectively, since each D_{2d} or S_4 minimum has eight or four symmetry related paths available corresponding

Table 2. Rate constants for rearrangements of two-dimensional LJ_7 calculated by dynamic path sampling [9] (DBC) and in the present work for $kT/\epsilon = 0.05$. Also tabulated is the ratio of the two calculated rate constants, where available, and a brief description of how the total rate decomposes into discrete path contributions.

Process	DBC	Present work	Ratio	Comment
$C_0^0 \rightarrow C_1^4$	3.50×10^{-13}	2.0×10^{-12}	0.18	86% from a one-step path
$C_0^0 \quad C_1^4$	7.25×10^{-4}	9.8×10^{-5}	7.40	86% from a one-step path
$C_0^1 \rightarrow C_2^1$	2.77×10^{-13}	3.6×10^{-13}	0.77	80% from a one-step path
$C_0^1 \quad C_2^1$	1.38×10^{-3}	1.7×10^{-4}	8.12	80% from a one-step path
$C_1^4 \rightarrow C_2^1$	5.30×10^{-7}	1.4×10^{-6}	0.38	56% from a one-step path
$C_1^4 \quad C_2^1$	1.39×10^{-6}	2.3×10^{-6}	0.60	56% from a one-step path
$C_1^4 \rightarrow C_1^1$	1.20×10^{-6}	2.4×10^{-6}	0.50	61% from a one-step path
$C_1^4 \rightarrow C_1^3$	5.80×10^{-7}	3.7×10^{-7}	1.57	50% from 3 three-step paths
$C_0^0 \rightarrow C_0^1$		6.0×10^{-14}		58% from 4 three-step paths
$C_0^0 \quad C_0^1$		1.0×10^{-14}		58% from 4 three-step paths

Table 3. Convergence of the microcanonical rate constants (in s^{-1}) for isomerization of $(H_2O)_8$ at $E = -253.13$ kJ/mol. Most of the contribution comes from the four-step path shown in figure 2.

Steps	Number of paths	$k_{D_{2d} \; S_4}$	$k_{S_4 \; D_{2d}}$
1	0	0	0
2	1	1.99×10^{-22}	2.09×10^{-22}
3	233	4.63×10^{3}	4.86×10^{3}
4	8929	1.09×10^{5}	1.14×10^{5}
5	238 314	1.11×10^{5}	1.16×10^{5}
6	6 834 382	1.11×10^{5}	1.16×10^{5}

to different permutational isomers of the C_1 symmetry intermediates.

An existing database of stationary points was first exploited to study the convergence of $k_{S_4 \; D_{2d}}$ and $k_{D_{2d} \; S_4}$ as paths containing increasing numbers of steps are included (table 3). Although the number of paths increases rapidly with the number of allowed steps, the rates converge once four-step paths are included. In fact, 90% of the total contribution to $k_{S_4 \; D_{2d}}$ and $k_{D_{2d} \; S_4}$ comes from one particular four-step path (figure 2). The recrossing correction for this path increases the rates by 50% giving a contribution of $3.2 \times 10^4 \, s^{-1}$ to $k_{D_{2d} \; S_4}$. The mechanism can be described in terms of four sequential donor-acceptor exchange rearrangements [86]. A further 4% of the total contribution to $k_{S_4 \; D_{2d}}$ and $k_{D_{2d} \; S_4}$ comes from the three-step path shown in figure 3; the recrossing correction is negligible in this case and the contribution to $k_{D_{2d} \; S_4}$ is $1.3 \times 10^3 \, s^{-1}$. Another 4% of the total $k_{S_4 \; D_{2d}}$ and $k_{D_{2d} \; S_4}$ rates comes from a four-step path (not shown) that is a small variation on the three-step path in figure 3: the highest energy transition state is the same and the contribution to $k_{D_{2d} \; S_4}$ is $1.1 \times 10^3 \, s^{-1}$. The microcanonical equilibrium occupation probabilities for particular permutational isomers at $E = -253.13$ kJ/mol are $P_{S_4} = 0.512$, $P_{D_{2d}} = 0.486$ and $P_I = 0.002$ for the 8876 intermediate minima in the database. In accord with detailed balance the $k_{S_4 \; D_{2d}}$ rates are about 5% faster at this energy than the corresponding $k_{D_{2d} \; S_4}$ rate.

In previous dynamical path sampling calculations $k_{D_{2d} \; S_4}$ was reported as $5 \times 10^4 \, s^{-1}$. This result was obtained by considering interconversions corresponding to rearrangements in a single face of the octamer cube [16]. The four-step path shown in figure 2 cannot be described in this way; instead the structure of the free energy transition state obtained from dynamical sampling [16] appears to correspond to the highest energy transition state on the three-step path shown in figure 3. The mechanism, involving intermediate local minima and transition states with diagonal hydrogen bonds across faces of the cube, also seems to agree with the description by Laria *et al*. [16]. The transition states considered in discrete path sampling are, of course, defined as stationary points with one negative Hessian eigenvalue, and must be clearly distinguished from a free energy transition state. However, for a path such as the one in figure 3, with a well-defined potential energy maximum, we might reasonably expect the structure of the free energy transition state to resemble that point. Summing the contributions of the two fastest paths that involve this transition state gives a contribution to $k_{D_{2d} \; S_4}$ of $2.5 \times 10^3 \, s^{-1}$, which rises to $1.0 \times 10^4 \, s^{-1}$ when the pathways differing only by permutational isomerization are included. Precise agreement with the dynamical path sampling results is not expected, in view of the approximations made in calculating partition functions and minimum-to-minimum rate constants in the present work.

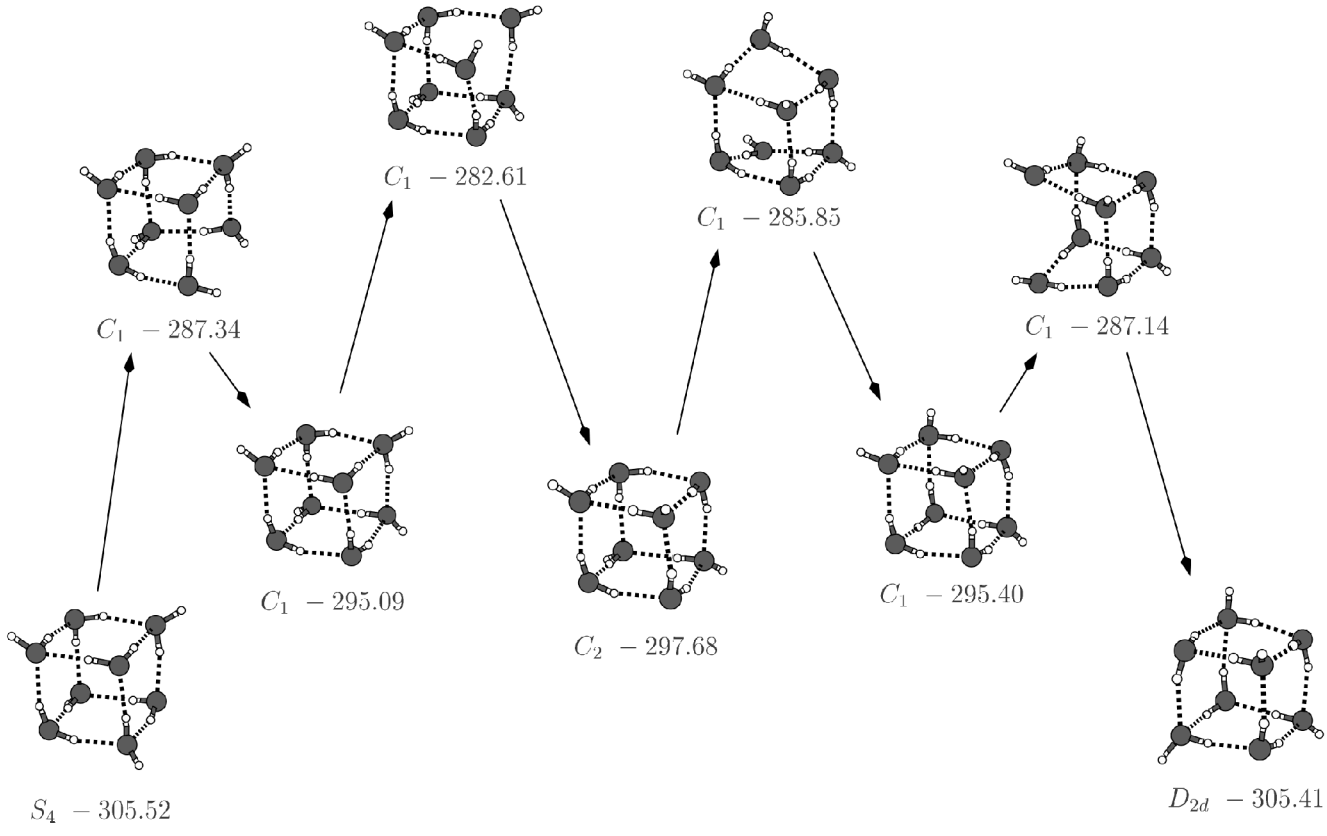

Figure 2. Four-step path (four transition states, five minima) linking S_4 and D_{2d} cuboids in the water octamer. The point group symmetries and energies (kJ/mol) for the TIP4P potential [91] are given below each stationary point.

A discrete path sampling calculation was then performed without using the existing database of stationary points that was employed in the above convergence tests. An initial four-step pathway was found by linking D_{2d} and S_4 minima, but with a negligible contribution to the total rate. The additional considerations required in treating a cluster of rigid body molecules are described in the Appendix. These calculations confirmed the pathways discussed above, with practically the same predicted rate constants. The three-step path becomes more important as the total energy increases, but is still calculated to be ten times slower than the fastest four-step path at $E = -240\,\epsilon$ within the current framework.

It would be interesting to know whether the path shown in figure 2 dominates the dynamics if it is admitted in a dynamical path sampling calculation. In fact, a standard molecular dynamics simulation of 10^{-6} s duration revealed pathways between the S_4 and D_{2d} minima, which appear to agree with the discrete path sampling results, although the statistics are not sufficient to provide converged rate constants [92]. In any case, the alternative paths based on structures containing diagonal hydrogen bonds are also easily located. The melting transition of this cluster would pose new problems, but cannot be described in terms of a simple rate constant, according to previous dynamical path sampling calculations [16].

7. Results for $(H_2O)_9$

Rodriguez *et al.* have recently reported a dynamical path sampling study of the isomerization rates between two low-energy isomers of $(H_2O)_9$ [20]. As for $(H_2O)_8$, above, the results are for the TIP4P rigid body intermolecular potential [91]. Comparison with the rate constant obtained from transition state theory in the harmonic approximation produced quantitative agreement for a one-step rearrangement between the global minimum of C_s symmetry (energy −344.436 kJ/mol) and a minimum with C_1 symmetry (energy −340.848 kJ/mol) [20]. The initial path calculated in the present work immediately reproduced this one-step mechanism, which is illustrated in figure 4 (top). The transition state involves an unusual three-centre hydrogen bond in which one hydrogen is coordinated to two different oxygens. The next-fastest path also involves a single transition state (figure 4 bottom), but the microcanonical rates are three orders of magnitude slower for

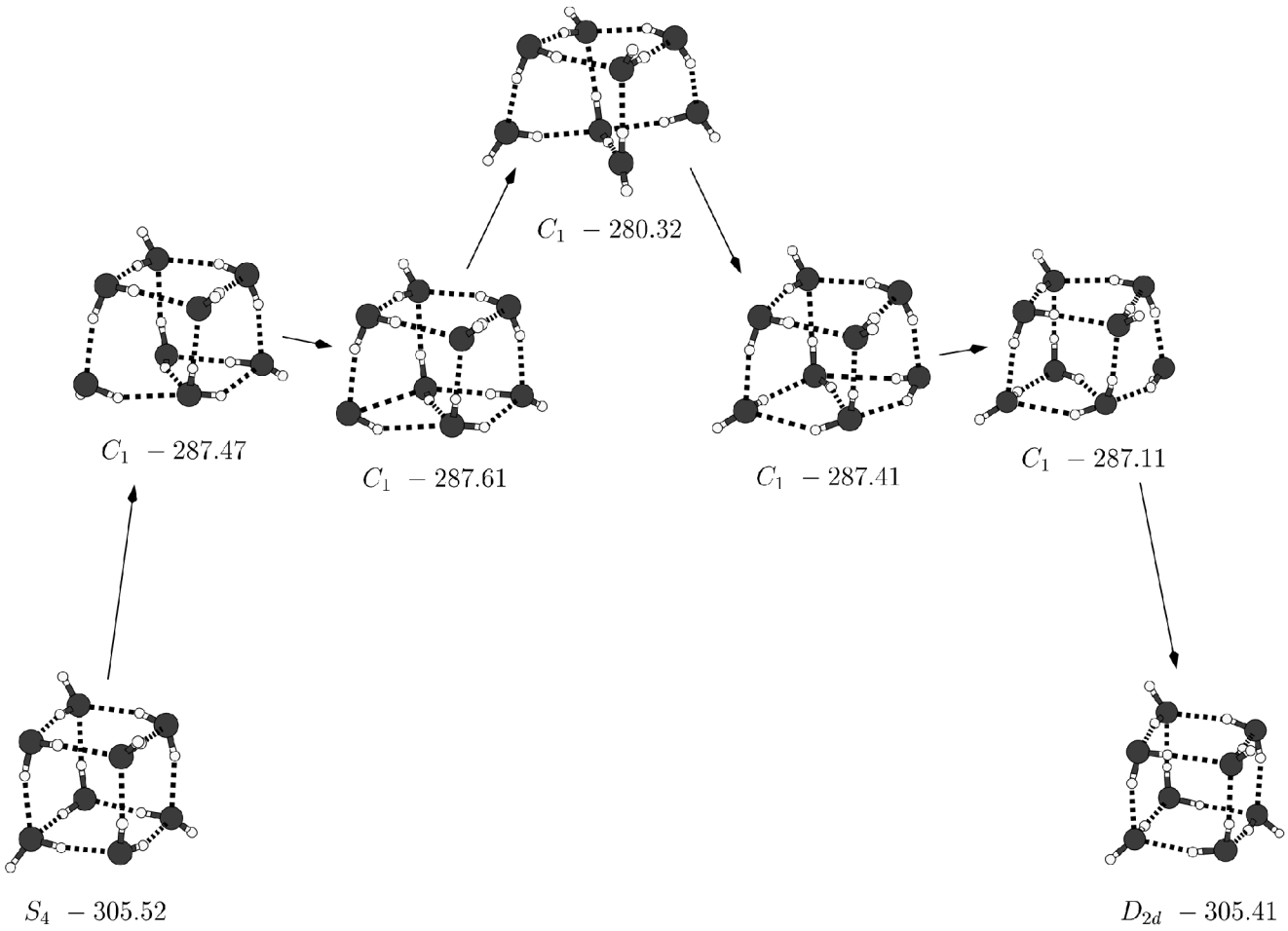

Figure 3. Three-step path (three transition states, four minima) linking S_4 and D_{2d} cuboids in the water octamer. The point group symmetries and energies (kJ/mol) for the TIP4P potential [91] are given below each stationary point.

the total energy considered by Rodriguez *et al.*, i.e. −317.984 kJ/mol. The higher energy transition state for this path has larger normal mode frequencies, which slows the rate by a factor of ten, while the energy difference accounts for the remaining factor of 100. A number of longer paths were also located, but their contribution to the isomerization rate at this total energy is negligible.

The slower rate constant involving the higher barrier is calculated as $4.1 \times 10^3\,s^{-1}$, while the faster rate is $8.9 \times 10^6\,s^{-1}$. These are the rate constants for isomerization between two specific permutational isomers. Every C_s symmetry permutational isomer of the global minimum is connected to two permutational isomers of the C_1 minimum, doubling the slower rate if permutational isomers are not discriminated. Rodriguez *et al.* report a value of $2 \times 10^3\,s^{-1}$ for the slower rate involving specific permutational isomers, both from transition state theory and dynamic path sampling. The difference between the rate constant calculated from transition state theory by Rodriguez *et al.* and our result is probably due to small discrepancies in the energies of the stationary points and their normal mode frequencies. For example, Rodriguez *et al.* report the energy and imaginary frequency of the transition state as −334.30 kJ/mol and $79.5i\,cm^{-1}$, respectively [20], whereas our values are −334.78 kJ/mol and $83.6i\,cm^{-1}$, which should be well converged. The different energy for the transition state alone would cause the calculated rate constant to change by a factor of three.

8. Results for LJ_{38}

The LJ_{38} cluster provides a useful benchmark system because the PES reflects the competition between two low energy morphologies, namely the global minimum truncated octahedron and incomplete icosahedra [93], which we will refer to as O_h and I_h for brevity. The stability balance between these structures is also important in various transition metal clusters [94–97]. The

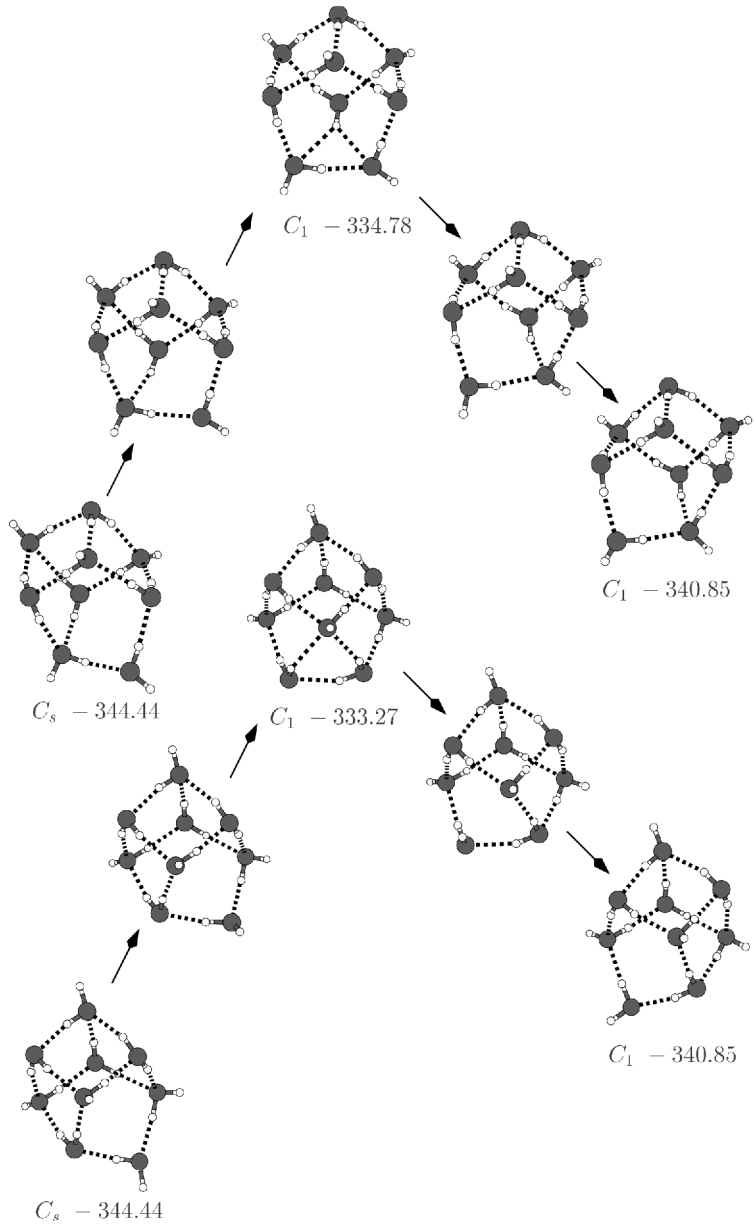

Figure 4. One-step paths for isomerization of $(H_2O)_9$; one intermediate geometry is included on each side of the transition state to clarify the mechanisms. The point group symmetries and energies (kJ/mol) for the TIP4P potential [91] are given below each stationary point.

heat capacity of this cluster exhibits a large peak at $kT/\epsilon \approx 0.18$, corresponding to the finite analogue of the solid-liquid melting transition [32, 98, 99]. However, there is a much smaller peak at $kT/\epsilon \approx 0.12$ corresponding to a solid-solid phase transition between low-lying minima based on truncated octahedra and incomplete icosahedra. The truncated octahedra are favoured in terms of potential energy, but icosahedra are favoured by entropy, since they are far more numerous, and generally have lower vibrational frequencies. The thermodynamics of this transition have previously been investigated by umbrella sampling [3] using a bond-orientational order parameter [99]. The free energy barrier at the equilibrium temperature was calculated to be $1.90\,\epsilon$, equivalent to about $17kT$ at a reduced temperature of $0.12\,\epsilon/k$.

Master equation dynamics for a sample of 6000 minima and 8633 transition states has previously been used to calculate $O_h \leftrightarrow I_h$ interconversion rates. An Arrhenius fit revealed an effective activation barrier that was close to the highest potential energy barrier for the lowest barrier path in the database, and

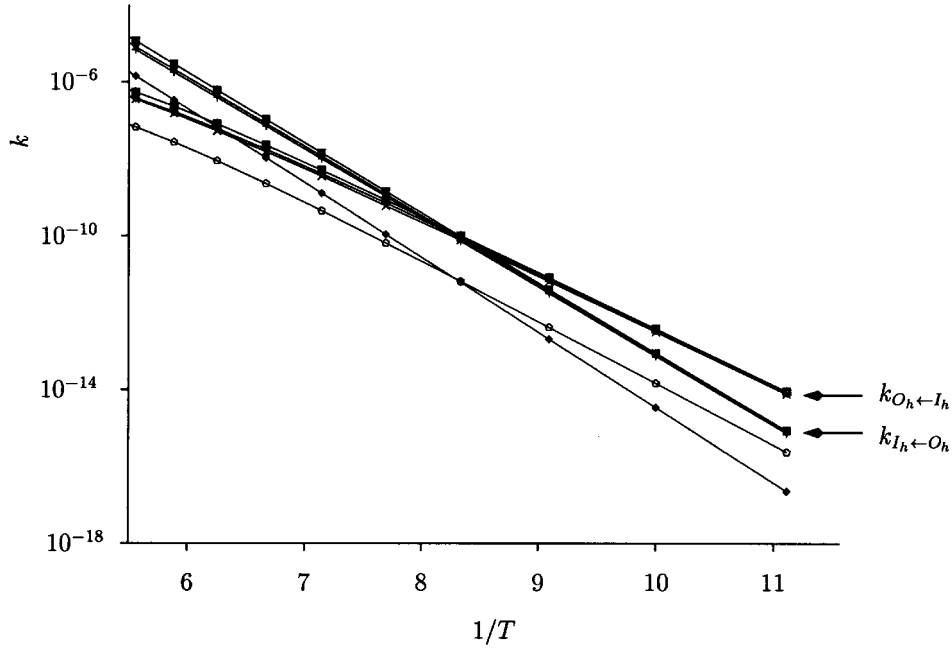

Figure 5. Arrhenius plots of $k_{O_h\ I_h}$ and $k_{I_h\ O_h}$ in LJ_{38} (reduced units). The closely superimposed curves are for the master equation results [32] and for cutoff parameter values of $\Delta = 0$, 10^{-9}, 10^{-8}, 10^{-7} and 10^{-6}. The shifted curves below are $k_{O_h\ I_h}$ and $k_{I_h\ O_h}$ for $\Delta = 10^{-5}$.

showed that the corresponding relaxation mode did indeed dominate the long time dynamics [99].

We first calculated $k_{I_h\ O_h}$ and $k_{O_h\ I_h}$ using the discrete path expressions in equation (8) for the same pathway database as was used in the previous master equation study. This exercise not only demonstrates that the discrete path formulation reproduces the master equation rates quantitatively, but also enables us to investigate how sensitive the calculation is to our assignment of local minima to the O_h, I_h and intermediate, I, states. This assignment could be done using the bond-orientational order parameters employed in previous work, but instead we simply considered the master equation eigenvector corresponding to the $O_h \leftrightarrow I_h$ interconversion. This eigenvector has one component, c_α, for each local minimum in the database, and the sign of the component enables us to classify the minima as I_h or O_h in character. The magnitude of the (normalised) components provides a systematic way of varying the number of intermediate minima by defining a cutoff value Δ. If $|c_\alpha| > \Delta$ then we classify minimum α as I_h or O_h, according to its sign, otherwise α belongs to the intermediate set. The results are shown in figure 5. The master equation rates and those calculated from equation (8) are practically identical for $\Delta = 0$, 10^{-9}, 10^{-8}, 10^{-7} and 10^{-6}, corresponding to 0% up to 60% intermediates (figure 5). However, for $\Delta = 10^{-5}$, where 77% of the local minima are classified as intermediates, the rates become significantly too small. Hence the rate expressions in equation (8) agree very well with the master equation rates so long as we do not try to apply the steady state approximation to local minima that are significantly populated in equilibrium.

We then proceeded to run true discrete path sampling calculations according to the scheme described in the Appendix, without using the existing databases of stationary points. The O_h and I_h regions were defined to contain five and 395 local minima, respectively, based on the assignment for $\Delta = 6 \times 10^{-6}$ applied to the database of 6000 minima described above. These sets reproduce the original master equation results to within 1% at low temperature and 25% at high temperature. Each run was started from a five-step connection that was found by a single double-ended search between randomly selected minima from the the O_h and I_h sets. The rate constants associated with this initial path make a negligible contribution to the overall rates.

Five discrete path sampling runs were performed initially to test the efficiency of different parameter choices, particularly the number of steps, n_{diff}, between the replacement minimum in the current path and the minima to which connections were then sought (see §4 and the Appendix). The five runs each terminated once fifty perturbations of the fifty fastest paths were performed, and the final $k_{I_h\ O_h}$ and $k_{O_h\ I_h}$ are practically the same in each case. The rates are enhanced by factors of about 1.4 and 3.7 over the previous master equation results, for low and high temperature, respectively (figure 6). The

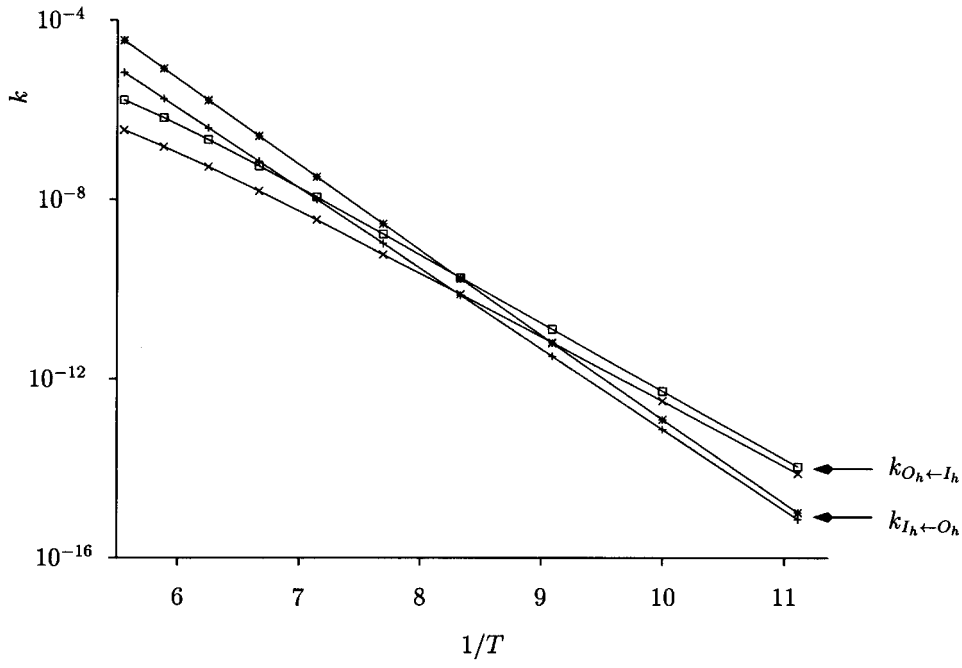

Figure 6. Arrhenius plots of $k_{O_h \ I_h}$ and $k_{I_h \ O_h}$ in LJ_{38} (reduced units) for the previous master equation results [32] and the present calculations using discrete path sampling. The new results correspond to the faster rate constants.

original five runs were all for a temperature of 0.09, at the lowest end of the range considered. However, we checked that the paths generated from a run at $kT/\epsilon = 0.12$ give essentially the same results. The fastest path contributes about 50% of the total rate at $kT/\epsilon = 0.09$; it involves four steps (figure 7) and the recrossing enhancement is 20%. In terms of the new paths located this fastest path was the 42nd found in the run with $n_{\text{diff}} = 4$, and the 207th found in the run with $n_{\text{diff}} = 2$. The ten fastest paths contribute more than 94% of the total rates in each case. When the corresponding stationary points are added to the original database used in previous work the resulting master equation rates become similar to the new discrete path sampling results.

Closer inspection of the pathway shown in figure 7 revealed that the second downhill path in fact exhibits a branch point: the true steepest-descent path involves a direct connection between two transition states [100, 101]. The second transition state also has C_s symmetry and mediates a symmetric degenerate rearrangement [63, 101] between permutational isomers of a C_1 minimum, as shown in the inset of figure 7. For a system with finite kinetic energy the branch point would be avoided and dynamical trajectories will relax into one of the two permutational isomers. Since the energy minimisation approach that we have employed to calculate pathways does not follow the steepest-descent path exactly, it also leads directly to a minimum. However, a plot of energy against path length reveals a distinct plateau as the optimisation passes around the second transition state. To fully account for the effect that such points might have on the dynamics would probably require more detailed dynamical calculations. Within the present framework we have simply allowed for the correct path degeneracy factor that results from the choice of paths leading to different permutational isomers.

Fitting the temperature variation of the discrete path sampling results to the Arrhenius form $A \exp(-E_a/kT)$ gives:

$$k_{I_h \ O_h}: \qquad A = 13.8\sqrt{\epsilon/m\sigma^2}, \quad E_a = 4.35\,\epsilon,$$
$$k_{O_h \ I_h}: \qquad A = 5.97\sqrt{\epsilon/m\sigma^2}, \quad E_a = 3.42\,\epsilon. \qquad (18)$$

Both the effective activation barriers and the prefactors are larger than for the original master equation results [32]. The effective barriers increase by about $0.2\,\epsilon$, and result in a steeper slope for the new results in figure 6. The two-step path also illustrated in figure 7 has a larger effective barrier. It contributes about 9% of the total rate at $kT/\epsilon = 0.09$, but becomes the second-fastest path at $kT/\epsilon = 0.12$ with a contribution of 20%. At this higher temperature the contribution of the four-step path in figure 7 falls significantly to about 27%. The fastest path found for this change of morphology in LJ_{38} is about 10^{14} times slower than the fastest iso-

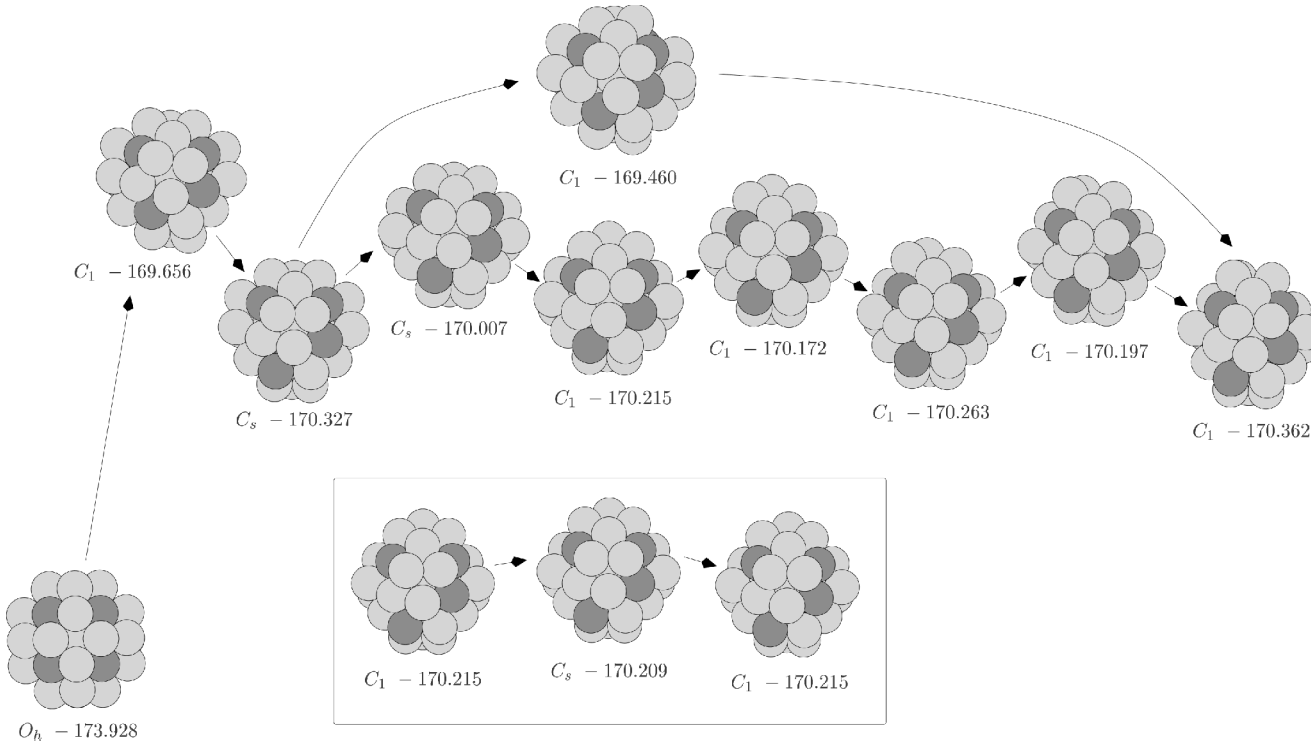

Figure 7. Four-step path (four transition states, five minima) linking members of the O_h and I_h sets in the LJ_{38} cluster. The symmetries and energies (ϵ) are given below each stationary point. The dark shaded atoms are the eight surface atoms that lie on C_3 rotation axes in the O_h minimum. The second downhill path involves a branch point as described in the text; the additional transition state has C_s symmetry and mediates a degenerate rearrangement of permutational isomers [63] as shown in the lower inset. A two-step path links different permutational isomers of the same two minima that are linked by the four-step path; the isomer shown for the final minimum is the one appropriate for the four-step path.

merization path described for two-dimensional LJ_7 in § 5 at a common temperature of $kT/\epsilon = 0.05$.

9. Conclusions

Previous calculations of dynamics and thermodynamics from databases of local minima and pathways have generally relied on sampling schemes that reflect the thermodynamic weight. For complex systems, with an exponentially large number of stationary points, dynamical criteria must be used to ensure proper sampling in the calculation of rate constants or relaxation times. The discrete path sampling theory, detailed in the present contribution, shows how the phenomenological rate constants for a two-state system can be written as a weighted sum of all the possible elementary paths. The resulting expressions are valid for a two-state equilibrium that is slow compared to all the other modes of relaxation, so long as the minimum to minimum dynamics are Markovian.

The theory necessitates calculation of the equilibrium populations of local minima, and also requires rate constants for the rearrangements mediated by a single transition state. In this initial study we have employed the simplest harmonic densities of states and transition state theory rate constants for these quantities. However, more sophisticated models could certainly be used, up to and including full molecular dynamics calculations using the discrete paths for guidance.

The relative efficiency of this discrete approach compared to fully dynamical path sampling is not yet known, and will likely be system dependent. However, the results in this initial report demonstrate that the discrete approach is, at least, formally correct. Agreement with previous dynamical path simulations, where available, is in line with the errors expected for harmonic transition state theory.

The LJ_{38} cluster exhibits competition between two morphologies. Rates calculated using the discrete path approach for the same database of stationary points agree with previous master equation results [99], so long as the number of local minima assigned as intermediates is not too high. Discrete path sampling runs for this system located faster pathways than the previous survey, which employed thermodynamic sampling criteria [99]. For parameters appropriate to argon the rate constants are about $80\,s^{-1}$ at the transition tempera-

ture. Since several distinct paths make significant contributions to this rate the LJ_{38} cluster may provide a useful testing ground for further path sampling studies, as it has already done for global optimisation [102, 103] and thermodynamics [32, 98, 99].

I am very grateful to Dr J. P. K. Doye and Dr F. Calvo for many helpful discussions, and to Dr M. A. Miller for providing the TIP4P $(H_2O)_8$ database.

Appendix

In this Appendix we provide details of the algorithms used for locating minimum-transition state-minimum $\cdots$ connections between pairs of local minima and for discrete path sampling.

Given the coordinates of two distinct local minima, which could be different permutational isomers, the discrete pathway search is outlined schematically in figure A1, and typical parameter values are given in table A1. As mentioned in § 4, double-ended transition state searches were conducted using the modified nudged elastic band (NEB) approach [25, 77–79] to generate a starting guess and hybrid eigenvector-following [32, 74, 75] to continue the search. The n_{neb} NEB images were initially equally spaced on the straight line path between the two minima in question, with a force constant of $|\Delta E|/100(n_{\mathrm{neb}}+1)\Delta s^2$, where ΔE and Δs are the energy difference and displacement between the two minima.

The basic strategy summarised in figure A1 is to seek successive transition states between pairs of minima in an ordered list, adding new minima if necessary. The shortest distance between two given structures was found by minimisation with respect to the three Euler angle displacements using analytic gradients [104] and a modified version of Liu and Nocedal's algorithm [73] (see § 4). The convergence criterion for these orientational optimisations was a root-mean-square gradient of less than 10^{-6}. When two minima (and an intervening transition state) were inserted in a path they were ordered using the shortest displacement to a neighbouring minimum in the list. The first few steps of a connection search involving well separated end points generally consist of adding new pairs of intervening minima in this way. Shortcuts, i.e. identical permutational isomers occurring at two places in the path, were identified as structures closer than $0.001\,\sigma$ after orientational optimisation. Minima were also removed from the list if it proved too difficult to connect them (figure A1); the neighbour of an end minimum was removed rather than that minimum itself.

The connection of pairs of local minima is an integral component of the overall discrete path sampling algorithm outlined in figure A2. Typical values of the additional parameters used in this part of the calculation are specified in table A2. As discussed in § 4, a randomly selected minimum in the current path is replaced either by a new minimum generated by perturbing the geometry and minimising, or by an off-pathway connected minimum that has already been located. Connections between the replacement minimum and the minima n_{shift} steps away from it in the path are then sought;

Table A1. Typical parameters used in the double-ended searches to connect pairs of local minima in discrete path sampling runs.

Parameter	Description	Value
$n_{\mathrm{cycle}}^{\mathrm{max}}$	maximum number of iterations	20
$n_{\mathrm{neb}}^{\mathrm{min}}$	minimum number of NEB images	5
$n_{\mathrm{neb}}^{\mathrm{max}}$	maximum number of NEB images	19
Δn_{neb}	increment in the number of NEB images	4
$n_{\mathrm{min}}^{\mathrm{max}}$	maximum number of minima in the path	101
$n_{\mathrm{ts}}^{\mathrm{max}}$	maximum number of transition states in the path	100
	maximum number of NEB iterations	50
	RMS force convergence criterion for NEB iterations	0.02
	container radius to prevent evaporation	$5\,\sigma$
	maximum step length in BFGS minimisations	$0.05\,\sigma$
	RMS force convergence criterion for geometry optimisation	$10^{-6}\,\epsilon/\sigma$
	maximum iterations for largest Hessian eigenvalue	200
	maximum iterations for smallest Hessian eigenvalue	200
	maximum iterations in tangent space near convergence	20
	maximum iterations in tangent space away from convergence	3
	percentage change for eigenvalue convergence	0.005
	pushoff in pathway calculations	$0.04\,\sigma$

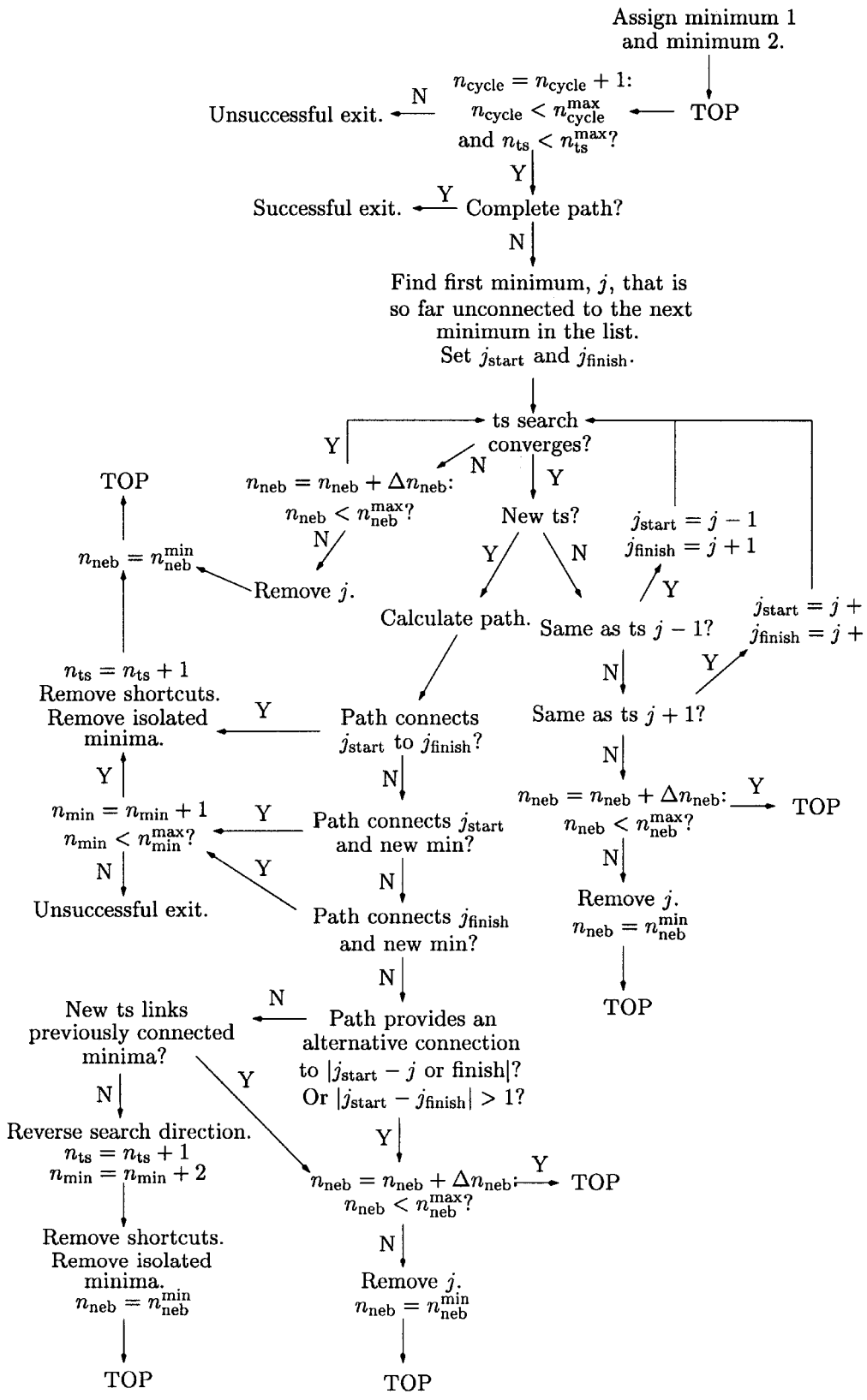

Figure A1. Schematic view of the algorithm used for double-ended searches to connect pairs of local minima via combined NEB and hybrid eigenvector-following calculations. Transition state j connects minima j and $j+1$ in the path. An isolated minimum is one with no forward or backward connection.

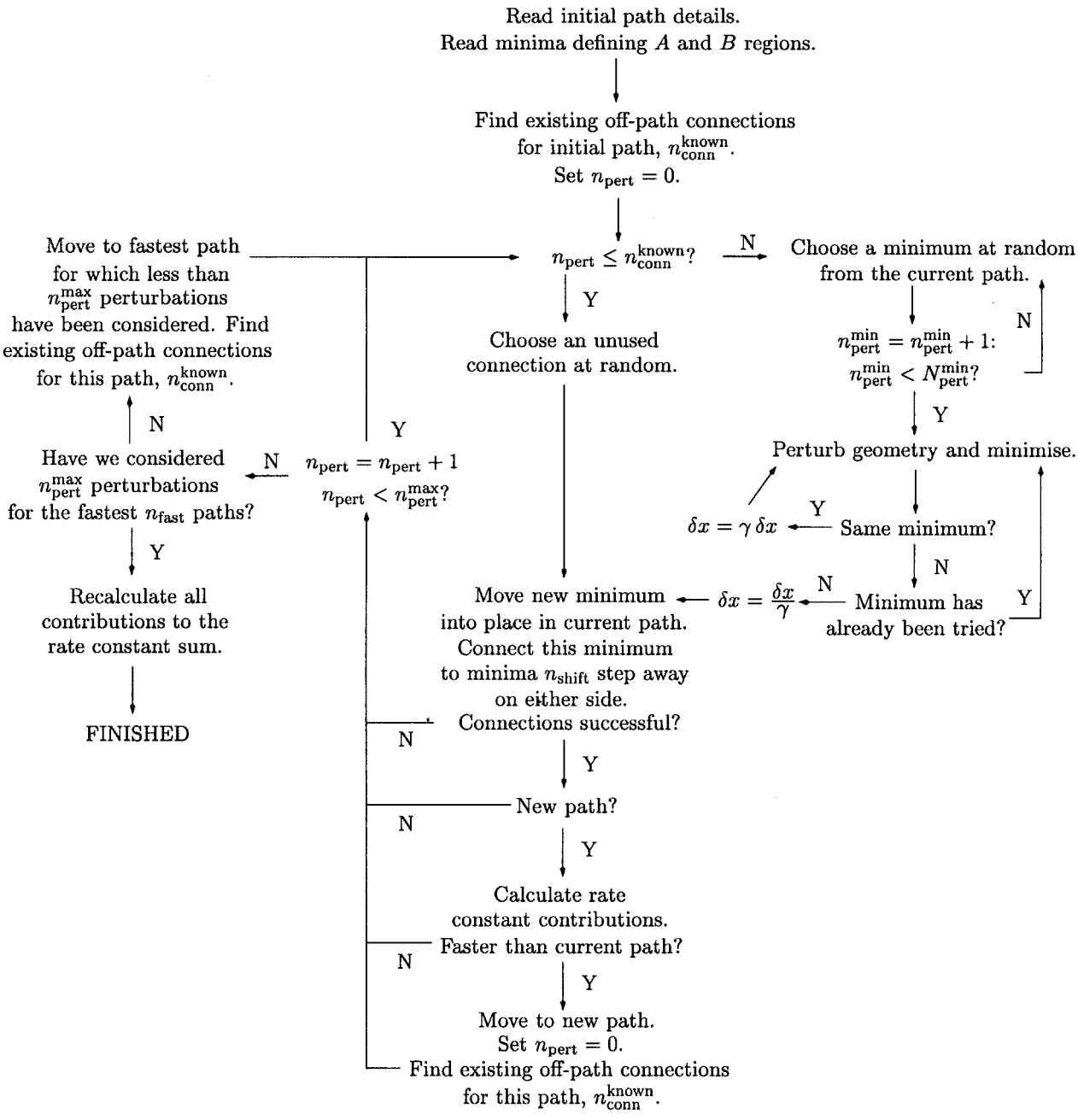

Figure A2. Schematic view of the discrete path-sampling algorithm.

the appropriate end minimum was used if n_{shift} steps lay outside the length of the path.

New minima, transition states and pathways were all recorded as they were found. For each new minimum additional single-ended transition state searches were run until a minimum number of connections had been achieved. Since further connections are continuously accumulated for the fastest pathways as the sampling run proceeds, this minimum value was chosen as three. The hard sphere collision theory used in previous work [76] was employed to generate initial guesses for these searches. The fraction of the first collision time was adjusted dynamically to ensure a good balance between the chance of finding a new transition state and the probability of it being connected to the initial minimum. When the rate constants were recalculated at the end of the run the connectivity of the minima in the fastest paths typically ranged from about 10 to 100.

It is important to allow for permutational isomerism. Equivalent stationary points were identified by com-

Table A2. Typical values of the additional parameters used in discrete path sampling runs, as described in the Appendix.

Parameter	Description	Value
$n_{\mathrm{pert}}^{\mathrm{max}}$	maximum number of perturbations per path	50
n_{fast}	$n_{\mathrm{pert}}^{\mathrm{max}}$ perturbations are performed for the n_{fast} fastest paths	50
$N_{\mathrm{pert}}^{\mathrm{min}}$	maximum number of perturbations for a local minimum	50
δx	maximum Cartesian coordinate displacement, initial value	0.9
γ	adjustment factor for δx	1.02
n_{shift}	connections are sought from a new minimum to minima n_{shift} steps away	1 to 4
	RMS force convergence criterion for 'sloppy' optimisations	$10^{-5}\,\epsilon/\sigma$
	smallest number of connections per minimum permitted	3
	maximum difference in energy for identical stationary points	$10^{-8}\,\epsilon$
	maximum difference in inertia tensor trace for identical stationary points	$0.01\,m\sigma^2$

paring both the energies and the trace of the inertia tensor evaluated for a coordinate system at the centre of mass; using the energy alone can result in ambiguity. Permutational isomers were then distinguished from identical isomers by calculating the minimum distance, as described above [104]. If permutational isomers are present in a path then a shorter and faster path can be made by removing the minima and transition states in between and applying an appropriate permutation-inversion operation to all the stationary points on one side of the connection. Each pathway was checked for this possibility.

To identify the permutation-inversion operation linking distinct isomers of a given stationary point we used the following approach, which employs orbits of the point group. An orbit is a minimal set of objects, in this case atoms, that map only amongst themselves under all the operations of the group. We first identify the orbit with the largest distance from the centre of mass, and move one of these atoms onto the z axis. We then find the atom with the largest distance from the z axis in the new orientation, and rotate it into the xz plane. If the first atom lies on a rotation axis of order h then there will be a choice of h such atoms. For the first isomer we can choose any of the members of this orbit. For the second isomer we try all the possibilities until we find the one where the orientation is the same, aside from permutations of labelled atoms. The permutation required to reorder the atoms of the second isomer to put them in correspondence with the first one is the required result. If none of the h orientations of the second isomer achieves this effect then a permutation-inversion operation must be required. In this case the coordinates of the second isomer are inverted through the centre of mass, and then we proceed as before. A suitable alternative procedure was developed for the two-dimensional cluster.

The above algorithm was used in two other parts of the discrete path sampling algorithm. Pathways were systematically checked to see if any of the transition states in the evolving database could link two of the minima and produce a shorter path. The permutational isomer of the saved transition state may not link the same permutational isomers that are in the path. We therefore recomputed the pathway for such transition states using a less stringent convergence criterion, simply to identify the precise permutational isomers of the minima that it linked. A shortcut pathway was then produced by finding the permutation-inversion needed to put one of the two minima in correspondence with the existing path. The remainder of the path was then put into correspondence with the second minimum. All such shortcut paths were recorded, along with their contributions to the rate constants, and were treated in the same way as the new paths located directly by perturbations. Of course, shorter paths that involve different transition states are not necessarily faster than the longer path.

The final consideration of permutational isomerism occurred when using saved transition states to make off-pathway connections to an adjacent minimum when perturbing the current path (see figure A2). Here again it is necessary to use the appropriate isomer of the transition state for the starting minimum, and the pathway was recalculated using weaker convergence criteria in order to establish this correspondence. These considerations of permutational isomerism may appear complicated, but they did not occupy a significant portion of the required computer time.

The calculations involving clusters of water molecules described as rigid bodies entailed additional bookkeeping. All the stationary point searches were conducted in centre-of-mass/Euler angle coordinates, as described in an earlier report [85, 86]. However, it was also necessary to transform to and from Cartesian coordinates to perform the minimum distance calculations

and characterise permutational isomers. In the NEB calculations the initial images were produced via interpolation in Cartesian coordinates, which does not preserve the rigid body geometries. These Cartesian coordinates were therefore transformed to best fit rigid body coordinates, and optimisation of the NEB images was performed in this coordinate system, keeping the Euler angles within the range $[0, \pi]$ or $[0, 2\pi]$, as appropriate. Pathways involving the exchange of two hydrogens on the same oxygen are quite common. This condition is easily diagnosed and was treated separately to avoid overlapping atoms in the initial Cartesian interpolation. The initial central image in this case was obtained by rotating the appropriate water molecule through an angle of $\pi/2$ about the C_2 symmetry axis.

References

[1] Bennett, C. H., 1997, in *Algorithms for Chemical Computations*, edited by R. E. Christofferson, American Chemical Society, Washington, D.C. (1997).
[2] Chandler, D., 1978, *J. chem. Phys.*, **68,** 2959.
[3] Torrie, G. M., and Valleau, J. P., 1974, *Chem. Phys. Lett.*, **28,** 578.
[4] Carter, E. A., Ciccotti, G., Hynes, J. T., and Kapral, R., 1989, *Chem. Phys. Lett.*, **156,** 472.
[5] Ciccotti, G., 1991, in *Computer Simulations in Materials Science*, edited by M. Meyer and V. Pontikis, pp. 365–396, Kluwer, Dordrecht.
[6] Rein ten Wolde, P., Ruiz-Montero, M. J., and Frenkel, D., 1996, *J. chem. Phys.*, **104,** 9932.
[7] Pratt, L. R., 1986, *J. chem. Phys.*, **85,** 5045.
[8] Dellago, C., Bolhuis, P. G., Csajka, F. S., and Chandler, D., 1998, *J. chem. Phys.*, **108,** 1964.
[9] Dellago, C., Bolhuis, P. G., and Chandler, D., 1998, *J. chem. Phys.*, **108,** 9236.
[10] Dellago, C., Bolhuis, P. G., and Chandler, D., 1999, *J. chem. Phys.*, **110,** 6617.
[11] Geissler, P. L., Dellago, C., and Chandler, D., 1999, *J. phys. Chem.* B, **103,** 3706.
[12] Geissler, P. L., Dellago, C., and Chandler, D., 1999, *Phys. Chem. Phys.*, **1,** 1317.
[13] Zuckerman, D. M., and Woolf, T. B., 1999, *J. chem. Phys.*, **111,** 9475.
[14] Bolhuis, P. G., Dellago, C., Geissler, P. L., and Chandler, D., 2000, *J. phys. Condensed Matter.*, **12,** A147.
[15] Marti, J., Csajka, F. S., and Chandler, D., 2000, *Chem. Phys. Lett.*, **328**, 169.
[16] Laria, D., Rodriguez, J., Dellago, C., and Chandler, D., 2001, *J. phys. Chem.* A, **105,** 2646.
[17] Crooks, G. E., and Chandler, D., 2001, *Phys. Rev.* E, **64,** 026109.
[18] Geissler, P. L., Dellago, C., Chandler, D., Hutter, J., and Parrinello, M., 2001, *Science*, **291,** 2121.
[19] Passerone, D., and Parrinello, M., 2001, *Phys. Rev. Lett.*, **87,** 108302.
[20] Rodriguez, J., Moriena, G., and Laria, D., 2002, *Chem. Phys. Lett.*, **356,** 147.
[21] Melchionna, S., 2000, *Phys. Rev.* E, **62,** 8762.
[22] Voter, A. F., 1986, *Phys. Rev.* B., **34,** 6819.
[23] Fichthorn, K. A., and Weinberg, W. H., 1991, *J. chem. Phys.*, **95,** 1090.
[24] Snurr, R. Q., Bell, A. T., and Theodorou, D. N., 1994, *J. phys. Chem.*, **98,** 11948.
[25] Henkelman, G., and Jónsson, H., 2001, *J. chem. Phys.*, **115,** 9657.
[26] Voter, A. F., 1997, *J. chem. Phys.*, **106**, 4665.
[27] Voter, A. F., 1997, *Phys. Rev. Lett.*, **78,** 3908.
[28] Voter, A. F., 1998, *Phys. Rev.* B, **57,** 13985.
[29] Sorensen, M. R., and Voter, A. F., 2000, *J. chem. Phys.*, **112,** 9599.
[30] Andricioaei, I., Straub, J. E., and Voter, A. F., 2001, *J. chem. Phys.*, **114,** 6994.
[31] Murrell, J. N., and Laidler, K. J., 1968, *Trans. Faraday Soc.*, **64,** 371.
[32] Wales, D. J., Doye, J. P. K., Miller, M. A., Mortenson, P. N., and Walsh, T. R., 2000, *Adv. Chem. Phys.*, **115,** 1.
[33] Stillinger, F. H., and Weber, T. A., 1983, *J. phys. Chem.*, **87,** 2833.
[34] Stillinger, F. H., and Weber, T. A., 1984, *Science* **225,** 983.
[35] Stillinger, F. H., 1999, *Phys. Rev.* E, **59,** 48.
[36] Ferrenberg, A. M., and Swendsen, R. H., 1988, *Phys. Rev. Lett.*, **61,** 2635.
[37] McDonald, I. R., and Singer, K., 1967, *Discuss. Faraday Soc.*, **43,** 40.
[38] Ferrenberg, A. M., and Swendsen, R. H., 1989, *Phys. Rev. Lett.*, **63,** 1195.
[39] Wales, D. J., 1993, *Mol. Phys.*, **78,** 151.
[40] Doye, J. P. K., and Wales, D. J., 1998, *Phys. Rev. Lett.*, **80,** 1357.
[41] Geyer, C. J., 1991, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, American Statistical Association, New York.
[42] Calvo, F., and Doye, J. P. K., 2001, *Phys. Rev.* E, **63,** 010902(R).
[43] Huberman, B. A., and Kerszberg, M., 1985, *J. phys.* A: Math. Gen., **18,** L331.
[44] Leopold, P. E., Montal, M., and Onuchic, J. N., 1992, *Proc. Natl. Acad. Sci.* USA, **89,** 8721.
[45] Berry, R. S., and Breitengraser-Kunz, R., 1995, *Phys. Rev. Lett.*, **74,** 3951.
[46] Kunz, R. E., and Berry, R. S., 1995, *J. chem. Phys.*, **103,** 1904.
[47] Kunz, B., Berry, R. S., and Astakhova, T., 1996, Surface Review and Letters, **3,** 307.
[48] Wang, J., Onuchic, J. N., and Wolynes, P. G., 1996, *Phys. Rev. Lett.*, **76,** 4861.
[49] Doye, J. P. K., and Wales, D. J., 1996, *J. chem. Phys.*, **105,** 8428.
[50] Becker, O. M., and Karplus, M., 1997, *J. chem. Phys.*, **106,** 1495.
[51] Kunz, R. E., Blaudeck, P., Hoffmann, K. H., and Berry, R. S., 1998, *J. chem. Phys.*, **108,** 2576.
[52] Walsh, T. R., and Wales, D. J., 1998, *J. chem. Phys.*, **109,** 6691.
[53] Cieplak, M., Henkel, M., Karbowski, J., and Banavar, J. R., 1998, *Phys. Rev. Lett.*, **80**, 3654.
[54] Ball, K. D., and Berry, R. S., 1998, *J. chem. Phys.*, **109,** 8541.
[55] Ball, K. D., and Berry, R. S., 1998, *J. chem. Phys.*, **109,** 8557.

[56] Kopsias, N. P., and Theodorou, D. N., 1998, *J. chem. Phys.*, **109,** 8573.
[57] Miller, M. A., and Wales, D. J., 1999, *J. chem. Phys.*, **111,** 6610.
[58] Westerberg. K. M., and Floudas, C. A., 1999, *J. chem. Phys.*, **110,** 9259.
[59] Ball, K. D., and Berry, R. S., 1999, *J. chem. Phys.*, **111,** 2060.
[60] Miller, M. A., Doye, J. P. K., and Wales, D. J., 1999, *J. chem. Phys.*, **110,** 328.
[61] Doye, J. P. K., and Wales, D. J., 1999, *Phys. Rev.* B, **59,** 2292.
[62] Miller, M. A., Doye, J. P. K., and Wales, D. J., 1999, *Phys. Rev.* E, **60,** 3701.
[63] Leone, R. E., and Schleyer, P. v. R., 1970, *Angew. Chem. Int. Ed. Engl.*, **9,** 860.
[64] van Kampen, N. G., 1981, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam.
[65] Kunz, R. E., and Berry, R. S., 1995, *J. chem. Phys.*, **103,** 1904.
[66] Mezey, P. G., 1987, *Potential energy hypersurfaces*, Elsevier, Amsterdam.
[67] Doye, J. P. K., and Wales, D. J., 2002, *J. chem. Phys.*, **116,** 3777.
[68] Chartrand, G., 1985, *Introductory graph theory*, Dover, New York.
[69] Horn, R. A., and Johnson, C. R., 1985, *Matrix Analysis*, Cambridge University Press, Cambridge.
[70] Chekmarev, S. F., and Krivov, S. V., 1998, *Phys. Rev.* E, **57,** 2445.
[71] Chekmarev, S. F., and Krivov, S. V., 1998, *Chem. Phys. Lett.*, **287,** 719.
[72] Calvo, F., Doye, J. P. K., and Wales, D. J., 2001, *J. chem. Phys.*, **114,** 7312.
[73] Liu, D., and Nocedal, J., 1989, *Mathematical Programming*, **45,** 503.
[74] Munro, L. J., and Wales, D. J., 1999, *Phys. Rev.* B, **59,** 3969.
[75] Kumeda, Y., Munro, L. J., and Wales, D. J., 2001, *Chem. Phys. Lett.*, **341,** 185.
[76] Middleton, T. F., and Wales, D. J., 2001, *Phys. Rev.* B, **64,** 024205.
[77] Henkelman, G., and Jónsson, H., 1999, *J. chem. Phys.*, **111,** 7010.
[78] Henkelman, G., Uberuaga, B. P., and Jónsson, H., 2000, *J. chem. Phys.*, **113,** 9901.
[79] Henkelman, G., and Jónsson, H., 2000, *J. chem. Phys.*, **113,** 9978.
[80] Liscomb, W. N., 1966, *Science*, **153,** 373.
[81] Mingos, D. M. P., and Wales, D. J., 1990, *Introduction to Cluster Chemistry*, Prentice-Hall (Englewood Cliffs).
[82] Tsai, C. J., and Jordan, K. D., 1991, *J. chem. Phys.*, **95,** 3850.
[83] Tsai, C. J., and Jordan, K. D., 1993, *J. chem. Phys.*, **99,** 6957.
[84] Tsai, C. J., and Jordan, K. D., 1993, *J. phys. Chem.*, **97,** 5208.
[85] Wales, D. J., and Ohmine, I., 1993, *J. chem. Phys.*, **98,** 7245.
[86] Wales, D. J., and Ohmine, I., 1993, *J. chem. Phys.*, **98,** 7257.
[87] Pedulla, J. M., and Jordan, K. D., 1998, *Chem. Phys.*, **239,** 593.
[88] Buck, U., Ettisher, I., Melzer, M., Buch, V., and Sadlej, J., 1998, *Phys. Rev. Lett.*, **80,** 2578.
[89] Gruenloh, C. J., Carney, J. R., Hagemeister, F. C., Arrington, C. A., Zwier, T. S., Fredericks, S. Y., Wood, J. T., and Jordan, K. D., 1998, *J. chem. Phys.*, **109,** 6601.
[90] Blanton, W. B., Gordonwylie, S. W., Clark, G. R., Jordan, K. D., Wood, J. T., Geiser, U., and Collins, T. J., 1999, *J. Am. Chem. Soc.*, **121,** 3551.
[91] Jorgensen, W. L., Chandraesekhar, J., Madura, J. W., Impey, R. W., and Klein, M. L., 1983, *J. chem. Phys.*, **79,** 926.
[92] Miller, M. A., personal communication.
[93] Wales, D. J., Miller, M. A., and Walsh, T. R., 1998, *Nature*, **394,** 758.
[94] Schmid, G., 1992, *Chem. Rev.*, **92,** 1709.
[95] Pinto, A., Pennisi, A. R., Faraci, G., D'Agostino, G., Mobilo, S., and Boscherini, F., 1995, *Phys. Rev.* B, **51,** 5315.
[96] Whetten, R. L., Khoury, J. T., Alvarez, M. M., Murthy, S., Vezmar, I., Wang, Z. L., Stephens, P. W., Cleveland, C. L., Luedtke, W. D., and Landman, U., 1996, *Advanced Materials*, **8**, 428.
[97] Andres, R. P., Bielefeld, J. D., Henderson, J. I., Janes, D. B., Kolagunta, V. R., Kubiak, C. P., Mahoney, W. J., and Osifchin, R. G., 1996, *Science*, **273,** 1690.
[98] Doye, J. P. K., Wales, D. J., and Miller, M. A., 1998, *J. chem. Phys.*, **109,** 8143.
[99] Doye, J. P. K., Miller, M. A., and Wales, D. J., 1999, *J. chem. Phys.*, **111,** 8417.
[100] Valtazanos, P., and Ruedenburg, K., 1986, *Theo. Chim. Acta*, **69,** 281.
[101] Wales, D. J., 2001, *Science*, **293,** 2067.
[102] Wales, D. J., and Doye, J. P. K., 1997, *J. phys. Chem.* A, **101,** 5111.
[103] Wales, D. J., and Scheraga, H. A., 1999, *Science*, **285,** 1368.
[104] Rhee, Y. M., 2000, *J. chem. Phys.*, **113,** 6021.