Conformational Similarities in Isomerization Dynamics of Clusters

Ersin Yurtsever*

College of Arts and Sciences, Koç University Rumelifeneri yolu, Sarıyer Istanbul, Turkey 34450

Ahmet Palazoğlu† and Yaman Arkun

College of Engineering, Koç University Rumelifeneri yolu, Sarıyer Istanbul, Turkey 34450

Received: April 25, 2003; In Final Form: June 5, 2003

A method for characterization of the isomerization dynamics from classical trajectories is presented. A measure function describing the topological distance between two clusters of atoms is first developed. Next, this measure is used to identify the regions of the potential energy surface visited by the trajectories. Unlike the commonly used techniques such as simulated annealing or quenching, the proposed method does not require repeated treatment of the trajectory and can be safely used to study the isomerization dynamics of large systems, especially those of monatomic clusters.

Introduction

The classical simulation methods provide a rich source of information for the thermodynamics and dynamics of many-body systems where the proper quantum mechanics is computationally prohibitive. In cluster research, especially for noble gas clusters, classical approaches have been used heavily to study various aspects of these interesting systems which form a bridge between the gas and bulk phase molecules. The topology of the potential energy surface (PES) in terms of the number and location of the minima and the first-order saddle points, the connectivity of these optimum points, the existence and the characteristics of the phase transitions, and finally the information on the reaction rates are some well-known examples of such studies.[1−8] Similarly, understanding the topology of the PES forms the basis of the structure and kinetics of conformational transitions of peptides leading to the protein folding problem. The characterization of the PES through its basins and the connectivity charts among them should help finding the reaction coordinate describing the folding pattern.[9−11]

Isomerization dynamics of clusters in principle can be analyzed by molecular dynamics simulations. For a fully ergodic trajectory, all of the regions of the phase space are sampled provided that the integration time is long enough. On the other hand, especially at low energy regimes, trajectories are not ergodic, and they sample the phase space in a quasiergodic manner.[12−13] In all cases, understanding of the dynamics requires information about the optimum points visited by the trajectory, the frequency of such visits, and the time spent around these regions. It is usually not so difficult to carry out this analysis for the minima, whereas measuring the time spent around the transition states (first-order saddle points) is more difficult as the curvature around these structures is generally high. One exception is the very shallow saddle of linear LJ_3, where the trajectories spend a considerable amount of time resulting in a relatively high stability described by a small Lyapunov exponent.[14−15]

The analysis of the frequency of visiting different minima as well as the time spent in the vicinity of these regions has been traditionally performed by two approaches: simulated annealing (SA) and quenching. In SA, the momentum of each particle is scaled such that the internal vibrational temperature is set almost to zero and then the new trajectory is integrated again with successive scaling until the particle is located in one of the minima of the PES. By proper optimization of the frequency of sampling and the scaling, the characterization can be completed. A more accurate method belongs to the class of optimization techniques such as steepest descent, conjugate gradient, and Newton and quasi-Newton methods, resulting in the so-called "quenching" approach.[12,16]

Both SA and quenching could provide reliable results in understanding the isomerization dynamics. However, quenching relies on the assumption that the local curvature of the PES should drive the cluster to the correct minima. This is a strong assumption especially for those cases where the trajectory is far from a specific minimum or may be equally "distant" to several minima at the same time. A more informative description of the trajectory should include not only the optimum points visited but also how close it gets to these points. Furthermore, any proposed method should preferably not require repeated computation of the trajectory as in optimization but rely on the information captured in the current state of the trajectory.

In this work, we present a method that quantifies the topological "distance" between various regions of the configuration-dependent part of the phase space. These distances can be cast into a measure of similarity. By using the similarity index, it becomes possible to quantify the distance from a single trajectory or a bundle of trajectories to the optimum points of the PES. This analysis should identify the visited regions of the phase space to give (a) the ergodicity of the trajectory and (b) the isomerizations taking place. The proposed method, unlike quenching, does not distort the trajectory to find the basin that it belongs to. Additionally, it is able to distinguish the time segments where the trajectory lies in relation to different basins contributing to the dynamics.

* To whom all correspondence must be addressed.

† On sabbatical leave from Department of Chemical Engineering and Materials Science, University of California, One Shields Avenue, Davis, CA 95616.

10.1021/jp035119j CCC: $25.00
Published on Web 07/12/2003

Methodology

The goal is to find an efficient and a robust way of evaluating the similarity (or dissimilarity) of the configurations expressed by two clusters of atoms. One can view these configurations as polyhedra whose vertexes correspond to the atoms. We propose to quantify the similarity measure between two cluster configurations by the distance between respective atoms. To calculate this measure in a meaningful way, the coordinate positions of the clusters need to be normalized, and the appropriate pairs need to be selected for comparison. In the following, we will introduce and discuss the steps of a normalization procedure for comparing shapes and then establish a similarity index (S).

The first step is to make the similarity measure invariant to the center of mass of the cluster. This is accomplished by defining all of the coordinates of particles in the center of mass coordinate system. Consider the following matrix of x,y,z-coordinate positions for 5 atoms in a cluster:

$$Z_c = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_4 & y_4 & z_4 \\ x_5 & y_5 & z_5 \end{bmatrix} \tag{1}$$

This matrix represents the spatial position of the cluster and uniquely defines its shape (configuration) centered on the point (0, 0, 0).

Singular Value Decomposition. When evaluating the geometrical similarity of two shapes, it is important to consider the possibility that the two shapes may be identical but oriented along different directions (rotated) or have different sizes (contracted or expanded). This would complicate the estimation of the similarity. To ensure that the similarity index is invariant to rotation and such scalings, as a second step in our normalization procedure, we perform singular value decomposition (SVD) of the matrix of coordinate positions to compute *the principal directions* along which the shape orients itself. In other words, we intend to align the matrix Z_c along its natural coordinate axes and not necessarily along the original x, y, and z axes.

Assume that Z_c is an lxm matrix (e.g., above, $l = 5$ and $m = 3$). Then, SVD of Z_c is performed as follows:

$$Z_c = U\Sigma V^{\mathrm{T}} \tag{2}$$

where U and V are unitary matrixes of size $l \times l$ and $m \times m$, respectively. The column vectors of U, denoted by u_i, are called the left singular vectors and correspond to the eigenvectors of the matrix $Z_cZ_c^{\mathrm{T}}$. On the other hand, the column vectors of V, denoted υ_i, are called the right singular vectors and correspond to the eigenvectors of the matrix $Z_c^{\mathrm{T}}Z_c$.[17] The matrix $\sum$ contains the nonnegative singular values, σ_i, arranged in descending order as in

$$\Sigma = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix}; \text{ if } l \geq m \text{ (our case)} \tag{3}$$

or

$$\Sigma = [\Sigma_1\ 0]; \text{ if } l \leq m \tag{4}$$

where $\sum_1 = diag\{\sigma_1,\sigma_2,....\sigma_k\}$ with $k = \min(l,m)$ and $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_k$. Equation 2 can also be expressed as follows:

$$Z_c = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i \upsilon_i^T \tag{5}$$

where r is the rank of Z_c.

The first right singular vector υ_1 represents the most dominant direction (largest variation) in the data (as it corresponds to the direction with the highest singular value). Therefore, υ_1 constitutes the first principal direction. Subsequent right singular vectors yield the other principal directions with decreasing importance. After the principal directions of a cluster are found, we need to compute how the clusters are distributed in each of these directions. We obtain this information by projecting the shape matrix Z_c on the principal directions (υ_i). For example, the ith projection is given by

$$Z_c\upsilon_i = \sigma_i u_i \tag{6}$$

when this projection is scaled by the ith singular value, one recovers the left singular vector u_i, which represents the contribution or "score" of the matrix Z_c to the ith principal direction

$$u_i = Z_c\upsilon_i/\sigma_i \tag{7}$$

Thus, the configuration matrix Z_c can be expressed in new coordinates as follows:

$$U = Z_cV\Sigma^{\#} \tag{8}$$

where # denotes the pseudoinverse of a matrix (i.e. $A^{\#} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$). By considering U (noting that only the first three columns of U are selected as the rank of $\sum$ is three), each configuration matrix is expressed in coordinates (principal directions) that are natural to that configuration and thus invariant to rotation. This ensures that the shape (configuration) is preserved in its "canonic" pose.

This analysis is strongly analogous to the Aquilanti's treatment of the quantum dynamics of three- and four-body systems in terms of the hyperspherical coordinates.[18−20] The partitioning of the phase space follows again a singular value decomposition of the coordinate matrix in the center of mass coordinate system. The singular values of $\sum$ are called kinematic invariants, and they are directly related to the moments of inertia of the system. They can also be cast into the form of a hyperradius and two deformations, which could be used to identify structural changes. The partitioning of the kinetic energy within this frame is also applied to the study of the dynamics of small clusters.[21]

Flipping. This step ensures proper comparison of objects that are mirror images of each other, i.e., the *chirality* problem.[22] The mirror images have equivalent intrinsic shape features but cannot be superimposed by translation and rotation. Although the previous step aligns the cluster configuration along the most dominant direction, it needs to be realized that the mirror image of the configuration along the principal axes is also an equally valid projection (has the same normalized moment values). Flipping is performed so that if two objects are mirror images of each other then the similarity measure can be calculated correctly. In the 3D object comparison literature, flipping entails counting the number of triplets on each side of the body and flipping the body such that the larger count is on the positive side.[23] In our application, we count the number of atoms on each side of the cluster (with respect to each coordinate axis) and flip the cluster so that the larger number of atoms appears on the positive side. In principle, we place the "larger" part of the cluster on the positive side.

The necessity of this operation can be understood when one considers a similar projection by diagonalizing the moment of inertia matrix. The eigenvalues give the length of each principal axis, and the eigenvectors define their direction. These eigenvectors can also be used to align clusters. However, their direction can be obtained in two equivalent ways, as the diagonalization procedure will choose randomly from two opposite directions. The singular value decomposition also suffers from this ambiguity; hence, the flipping should be performed. Although it is also possible to achieve this procedure by trying out 8 possible orientations of three eigenvectors ($+++$, $++-$, $+-+$, $-++$, $--+$, $-+-$, $+--$, $---$), it is obvious that flipping is superior. Trying all possible orientations requires completion of the algorithm to find the similarity index whereas flipping is a single-step procedure and independent of the steps described below.

Similarity Index. The measure of similarity used in this work is based on the root-mean-square (RMS) distance calculated between the atoms from each cluster. The coordinate positions associated with each atom in each cluster are those that result after the preceding centering, SVD, and flipping steps. Specifically, if there are N atoms present in a cluster and if x_i^A, y_i^A, z_i^A denote the position of an atom i in the cluster A and x_j^B, y_j^B, z_j^B denote the position of an atom j in the cluster B, then the pairwise atomic distance can be expressed as

$$d_{ij} = \sqrt{(x_i^A - x_j^B)^2 + (y_i^A - y_j^B)^2 + (z_i^A - z_j^B)^2} \qquad (9)$$

and the similarity index between cluster A and cluster B is defined as

$$S_{AB} = \sum_{i,j=1}^{N} d_{ij} \qquad (10)$$

Thus, S_{AB} represents the sum of distances between N pairs, and a small value of S_{AB} indicates that the two clusters are geometrically similar. For molecules with specific backbone structures or clearly labeled positions, such measures are very easy to define and have been in use for a long time. However, because each atom is identical in a given monatomic cluster, how to select the relevant pairs from two clusters to be compared in eq 9 is not obvious. Next, we discuss an assignment technique to remove this ambiguity.

Assignment. This step solves the assignment problem in which the distance measure is calculated by optimal matching of the relevant atoms. Our aim is to find the pairs that yield the minimum similarity index (S) between two clusters. This turns out to be the well-known assignment problem and the algorithm to solve the problem is often referred as the "Hungarian method".[24−25] This formulation solves a linear programming problem to decide on the pairs of atoms that result in the minimum S.

$$S = \min_{a_{ij}} J = \sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij}a_{ij}$$

$$\sum_{i=1}^{N} a_{ij} = 1$$

$$\sum_{j=1}^{N} a_{ij} = 1$$

$$a_{ij} > 0 \qquad (11)$$

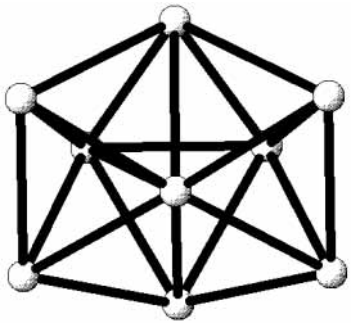

Figure 1. Global minimum of LJ_9.

Here, a_{ij} are the binary variables (0/1) that define if two atoms are matched (1) or not (0). d_{ij} defines the RMS distance between the ith and the jth atoms of two configurations that are being compared. The objective function J is the sum of distances between all matched atom pairs, defining a total distance measure between two configurations. S is then found by the minimum value of J.

An illustration of this method applied to a four-atom cluster is given in Appendix A.

Remark. Distance measures have been studied before in the context of constructing the PES. The main concern in establishing a distance measure has been the translation and rotation dependence of the compared structures. Translation invariance was simply achieved by translating the compared structures such that their centroids lie at the origin. To achieve rotational invariance, Thompson et al.[26] suggested a closed-form solution for this matrix, and Rhee[27] proposed an algorithm that can compute the rotational transform using an optimization procedure. Curotto et al.[28] generated a set of operations to help check for similar structures in constructing the PES. These studies do not tackle the problems of scaling, chirality, and the assignment of pairs, which we have addressed in our formulation.

Isomerization Dynamics of LJ_9

We have applied this analysis to the isomerization dynamics of a nine-atomic cluster held together by Lennard-Jones (LJ) interactions. In the following discussion, we use the reduced units where the mass of the particle, σ (hard-sphere diameter), and ϵ (interaction strength) of the LJ function are taken as unity and the other variables are scaled accordingly. The energy of the global minimum of the cluster is -24.113361 with C_{2v} symmetry (Figure 1) and PES has 20 distinct local minima (Table 1) excluding the permutational isomers. We have constructed our similarity index such that it does not distinguish these permutational isomers as it should not.

Trajectories start at the global minimum, with random momenta selected such that the total linear momentum, angular momentum, and their components are zero. We achieve that by randomly mixing the eigenvectors corresponding to normal modes of the starting structure and using them as momentum vectors. The magnitudes of the momentum vectors are scaled to desired energy. For various energy values, the directions of the atomic momentum vectors are the same so that similar motion at different energy can be compared. The Hamilton's equations of motion are solved with Runge−Kutta integration, which keeps the error in energy around 10^{-6}. Trajectories are integrated for 10^5 steps with each time step of 0.00125 time units. At five different total energy values, the snapshots are taken at every 50 steps and stored for assignments using quenching and similarity analyses. The characterization of the trajectory is done by using the database of minima given in Table 1.

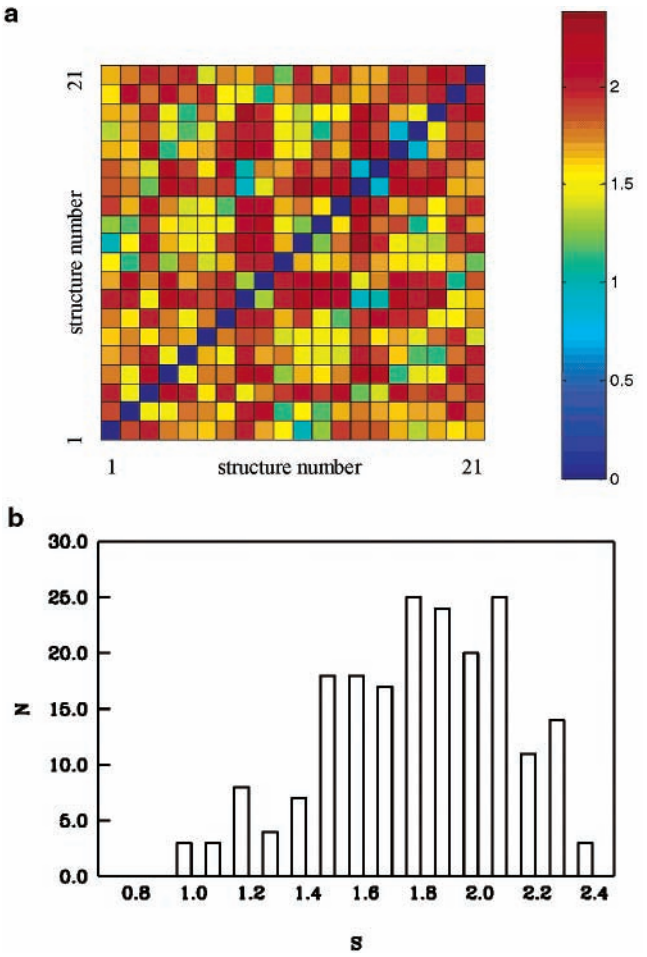

Figure 2. a. Contour plot of S for the minima of LJ_9. b. Histogram of S for the minima of LJ_9.

TABLE 1: Minima for LJ_9

structure	energy	symmetry group
1	−24.113361	C_{2v}
2	−23.269812	D_{3h}
3	−23.235647	C_{2v}
4	−23.196953	C_s
5	−23.173159	C_1
6	−23.170759	C_2
7	−23.149319	C_2
8	−23.105019	C_s
9	−23.043510	C_s
10	−23.000683	D_{3h}
11	−22.960888	C_s
12	−22.954098	C_s
13	−22.943932	C_1
14	−22.927790	C_{2v}
15	−22.843670	C_s
16	−22.474901	C_{3v}
17	−22.186206	C_s
18	−22.180733	C_2
19	−22.165974	C_1
20	−22.081303	C_1
21	−22.031475	C_2

Similarity Statistics of Minima. In Figure 2a, we present the distribution of S between all pairs of minima. The smallest values belong to the minima pairs of 14−15, 16−17, and 8−14 (S around 0.95). The global minimum (structure 1) is the closest (most similar) to the structure of minimum 11 with an S value of 1.00. In Figure 2b, we present the histogram for S between all minima. It is seen that the distribution affected by the similarity index provides a reasonable range for distinguishing various structures.

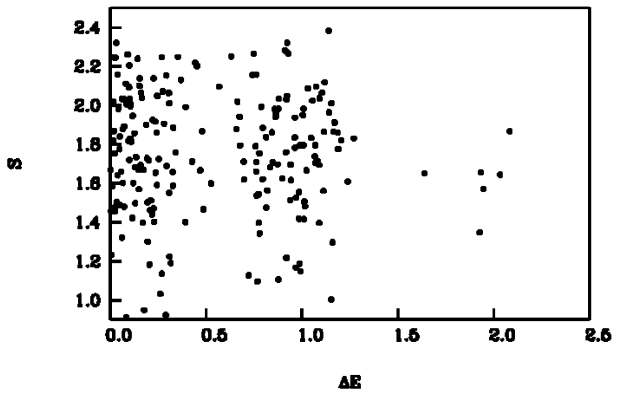

Figure 3. S vs ΔE for the minima of LJ_9.

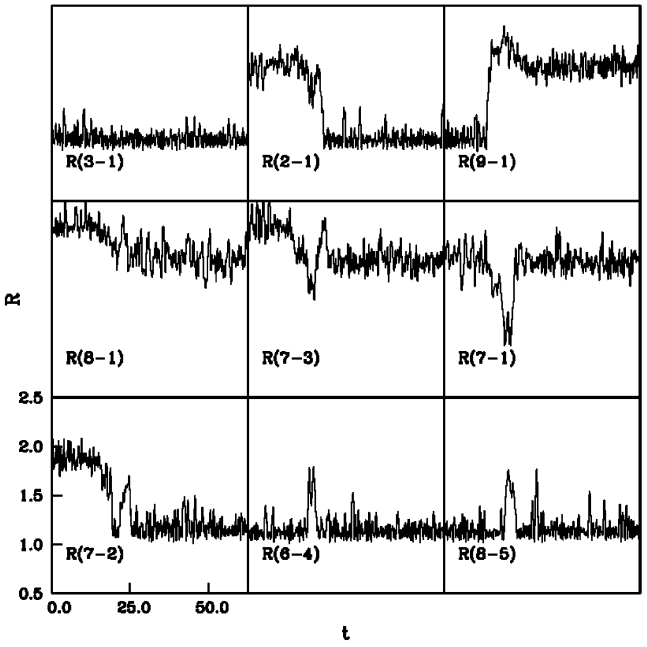

Figure 4. Bond order fluctuations at $E/\epsilon = -19.8$. (R in units of σ).

In Figure 3, the plot of S vs ΔE for all pairs of minimum structures is given. The value for S clearly does not depend on whether two minima are energetically close along the PES. For almost degenerate cases, S can be as large as 2.3 and could be as low as 0.9 for structures with large energy differences. Even though this is somewhat expected, it is ensuring that the index is not to be biased by energetics of the local minima.

Analysis of Bond Lengths. To follow the details of the isomerization of LJ_9, we have computed the bond lengths during the simulation. An analysis of the changes in the bond angles, as the simulation proceeds should give us additional insight into the structural phenomena that may be taking place. In the energy range $E > -19.8$, all of the bond distances fluctuate around their respective mean values. However, as seen in Figure 4, at this energy, sudden changes in the bond lengths are noticeable. Figure 4 depicts a set of representative bond lengths as a time-series to illustrate several different behaviors. One group of bonds (there are 16 of them in this group) oscillates around the mean similar to the low energy trajectories. R_{3-1} is given as an example of this group. Moreover, there are six bonds, which display sudden and large changes around $t = 25$. Two examples are R_{2-1} and R_{9-1}, where the changes in bond lengths are almost as large as the hard sphere diameter. The time period just short

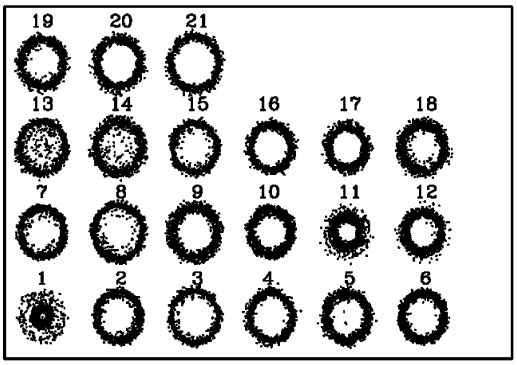

Figure 5. Visualization of the visits to minima by a single trajectory at $E/\epsilon = -19.8$.

of these "changes" reveals slightly different characteristics than the initial period, which exhibits almost harmonic oscillations. Finally, there is one more group of bonds that displays a great deal of activity around $t = 15$ (e.g., R_{2-1} and R_{2-3}) and then settles either at the initial bond length or at a slightly different one. All of these changes may imply strong isomerization taking place at this energy; however, the nature of these changes requires further analysis. We intend to provide a context for these behaviors using the similarity index, next.

Similarity Analysis of Dynamic Trajectories. The coordinate-dependent portion of the trajectory is described as a time series in $3N - 6$ dimensional phase space. A meaningful visualization of the minima visited by this time series is difficult and not unique, but we used a representation, which could be helpful for such purposes. In Figure 5, each minimum is represented by a point on a 2D grid. At each time step, the similarity index to a specific minimum (k) is defined by a random point on a circle with radius S_k. When the trajectory is integrated for a sufficiently long time, the distribution of the points should describe the "closeness" of the trajectory to the minimum (k). In this plot, a doughnut-shaped disk denotes a minimum that is avoided by the trajectory and an almost filled circle implies that the trajectory spends a considerable time in the vicinity of this minimum. In Figure 5, the minima are numbered starting from the lower left corner corresponding to the global minimum, and the numbering proceeds from left to right and to the higher rows. The energy is $E = -19.8$, and the starting conformation is the global minimum. It can be clearly seen that the trajectory spends most of its time around the global minimum. We also note that the minima 13 and 14 are highly visited as well as minimum 11. It is also notable that the trajectory does not stray away from 11; however, it could be very distant to 13 and 14. Moreover, most of the high-lying minima (e.g., minima 16, 17, 19, 20, and 21) are almost never visited.

To provide a comparison, we proceeded to quench structures obtained by sampling the history files. For this purpose, we used the program package OPTIM.2.3. The snapshots from trajectories are used as initial guesses, and they are optimized by the conjugated gradient method and the minima found in this manner are stored. The same initial structures of the quenching process are also subjected to our similarity analysis. Each of these structures is projected on all 21 minima by properly rotating, scaling, and assigning the corresponding pair of atoms, and then, the similarity index measuring the conformational distance to minima are calculated. The conformation with the minimum similarity index is chosen to characterize the trajectory at that time step.

In Figure 6, the results from a single trajectory at $E = -19.8$ are presented for comparison with those from quenching. The

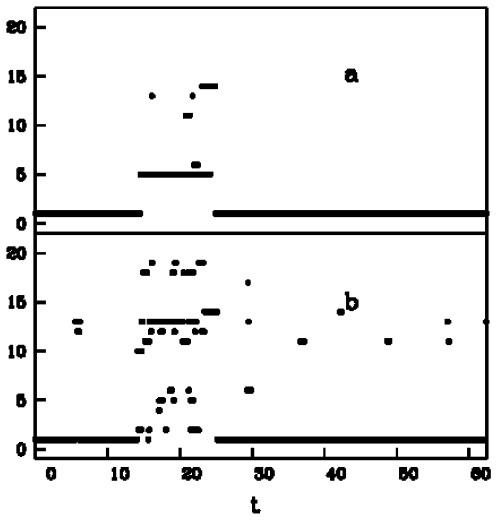

Figure 6. Minima visited by the trajectory as captured by (a) quenching and (b) similarity analysis.

x axis is the time or, alternatively, is the snapshots from time series, and the y axis is the label of the minima found from both methods defined as referenced in Table 1. For the method based on similarity analysis, we select the minimum that is the closest (most similar configuration) to that snapshot.

In Figure 6, there are three regions of interest, $t < 15$, $15 < t < 25$, and $t > 25$. During the initial stages of the simulation until $t = 15$, the trajectory seems to be staying in the vicinity of the global minimum, which can also be deduced from the bond lengths, the results of quenching as well as similarity analysis. The trajectory is also close to the global minimum after $t = 25$, but structural changes are suggested around $15 < t < 25$. This seems to be in contrast with Figure 4 where the drastic shifts in bond lengths occur in the third region ($t > 25$). A possible explanation for this discrepancy is that a permutational isomer of the global minimum is reached. That is, for $t > 25$, the cluster has a very similar shape to the global minimum as indicated by the similarity index; however, the assignments of the apex atoms have been changed as reflected in the bond lengths. [Actually, the detection of the permutational isomers is very easy in our approach; the similarity index without the assignment part should give a nonzero measure whereupon solving the assignment problem a zero measure should be obtained.] On the other hand, the geometrical isomerization takes place in the second region where a careful look also reveals more subtle changes in the last six plots of bond lengths. These changes are picked up by both quenching and the similarity analysis. However, it is notable that they characterize these structures as belonging to different minima. In quenching, the change is mainly to the structure 5, whereas the similarity index identifies a mixture of structures located around 13 (there are also few instances where 5 is found by similarity and 13 by quenching).

To test the reliability of our method, we compared the topological distances of the trajectory to those minima located by quenching and similarity analysis. As we selected the single minimum structure with the smallest similarity index (most similar configuration), one may argue that there might be more than one structure with slightly different S values, meaning that the configuration for that snapshot might be almost equidistant (equally similar) to the configurations of more than one minimum. In those cases, identification of the basin cannot be made with certainty. However, we note that the errors associated with such close-lying S values could easily be handled by

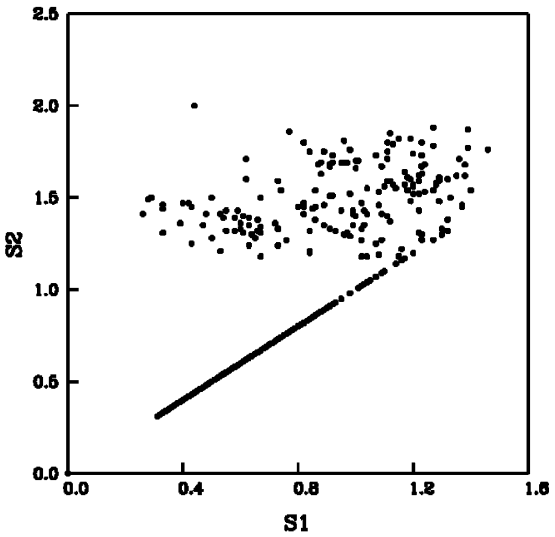

Figure 7. Comparison of minima captured by quenching and similarity analysis.

averaging over a bundle of trajectories, thus reducing the statistical errors.

In Figure 7, the similarity indices of the minima found by quenching are compared to those from our analysis. In other words, the plot shows the distance of the minima at a given point along the trajectory and compares it with the distance of the minima identified by quenching. When both methods find the same structure, the points should align themselves along the diagonal. On the other hand, if our method locates different structures with slightly different *S* values as we suspected, again the almost diagonal character of the plot should prevail. Indeed, for the trajectories studied at $E = -19.8$ and at a higher energy regime, many of these points are along the diagonal but there are also a large number of cases where the distance (similarity) to the minima found by quenching is very large (highly dissimilar). Both methods seem to characterize the sampled regions of the phase space differently. It is notable that when the trajectory is close to local minima both methods appear to agree reasonably well except at isolated points (e.g., when $S < 0.6$). Yet, when the trajectory strays away from the minima (as indicated by higher *S* values), the disagreement between the two methods is striking. Indeed, when the two methods disagree, the scatter of points is above the diagonal, indicating that quenching locates minima, which may be significantly farther (dissimilar) than the configuration of the cluster at that time step.

Conclusions

We have presented a robust method of identifying the sampling of the phase space in classical dynamical simulations. The method combines scaling, rotating, and finding similarity of structures in terms of distances in the conformational space. It can be used to study the quality of results by comparing structures of clusters obtained by different methods. It can provide a measure of the ergodicity of the simulation. However, most of all, it provides an efficient way to detect structural changes, isomerizations, during the simulations even for those cases where the bond lengths may not show significant activity. The basic advantage over quenching is that one does not have to choose a path among the many with negative eigenvalues to find a minimum; transitions and isomerizations can be followed in an automated manner.

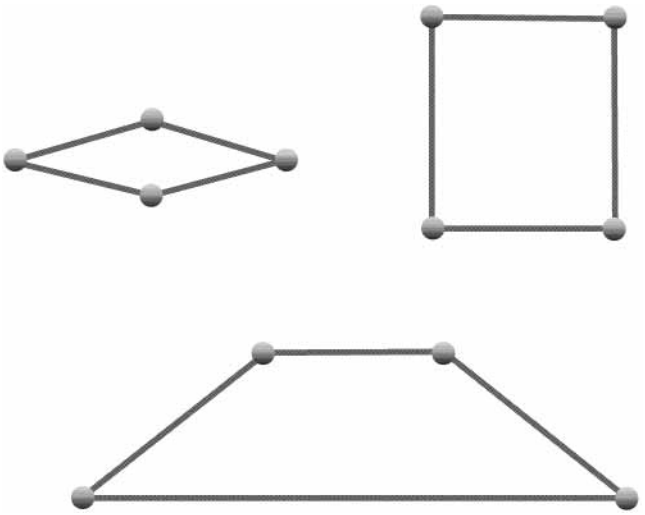

Figure 8. Starting shapes for comparison.

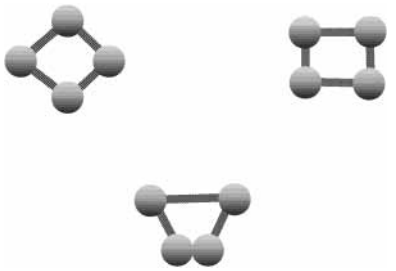

Figure 9. Shapes after normalization.

Acknowledgment. We would like to thank D. J. Wales for giving us a copy of OPTIM.2.3 and M. Turkay for pointing us to the assignment method. E.Y. also acknowledges the partial support of the Turkish Academy of Sciences (TÜBA).

Appendix A

The computation of the similarity index is illustrated for a planar test configuration, which represents a rhombus, given below:

$$Z_c = \begin{bmatrix} 2 & 2 \\ -2 & 2 \\ -6 & -2 \\ 6 & -2 \end{bmatrix} \tag{A1}$$

This configuration will be compared with two basic shapes, a diamond (*D*) and a rectangle (*R*)

$$D_c = \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ -3 & 0 \\ 0 & -1 \end{bmatrix}; R_c = \begin{bmatrix} 2 & 3 \\ -2 & 3 \\ -2 & -3 \\ 2 & -3 \end{bmatrix} \tag{A2}$$

Figure 8 depicts these configurations, Z_c, D_c, and R_c. Note that each configuration is in the center of mass reference frame.

The next step involves the computation of the SVD for each of these configurations to determine their natural alignment directions. For the test configuration, note that the matrix $Z_c^{\mathrm{T}}Z_c$ is given as

$$Z_c^{\mathrm{T}}Z_c = \begin{bmatrix} 80 & 0 \\ 0 & 16 \end{bmatrix} \tag{A3}$$

We can calculate that the singular values of Z_c are $\sigma_1 = 8.94$ and $\sigma_2 = 4.0$ and the eigenvectors for $Z_c^T Z_c$ are $v_1 = (-1\ 0)$ and $v_2 = (0\ -1)$. In the next step, each configuration is projected onto the directions indicated by their eigenvectors. The coordinates for each configuration are given as

$$Z_c V_{Zc} = \begin{bmatrix} -2 & -2 \\ 2 & -2 \\ 6 & 2 \\ -6 & 2 \end{bmatrix}; \quad D_c V_{Dc} = \begin{bmatrix} -3 & -2 \\ -3 & -2 \\ 3 & 2 \\ 3 & -2 \end{bmatrix};$$

$$R_c V_{Rc} = \begin{bmatrix} -3 & 0 \\ 0 & -1 \\ 3 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{(A4)}$$

Finally, the score matrices are obtained after scaling by the singular vectors. These are final coordinates to be compared. Figure 9 displays these configurations whose coordinates are given below:

$$U_{Zc} = \begin{bmatrix} -0.2236 & -0.5 \\ 0.2236 & -0.5 \\ 0.6708 & 0.5 \\ -0.6708 & 0.5 \end{bmatrix};$$

$$U_{Dc} = \begin{bmatrix} -0.7071 & 0 \\ 0 & -0.7071 \\ 0.7071 & 0 \\ 0 & 0.7071 \end{bmatrix};$$

$$U_{Rc} = \begin{bmatrix} -0.5 & -0.5 \\ -0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \quad \text{(A5)}$$

The similarity index is then calculated between the test configuration and the basic configurations as follows:

$$S_{Z-R} = 0.8944; \quad S_{Z-D} = 2.2037$$

This shows that the given rhombus is actually more similar to the rectangular than the diamond configuration.

References and Notes

(1) Berry, R. S. *J. Phys. Chem.* **1994**, *98*, 6910.
(2) Doye, J. P. K.; Wales, D. J. *Science* **1996**, *271*, 484.
(3) Wales, D. J. *Science* **1996**, *271*, 925.
(4) Komatsuzaki, T.; Berry, R. S. *J. Chem. Phys.* **1999**, *110*, 9160.
(5) Kunz, R. E.; Berry, R. S. *J. Chem. Phys.* **1995**, *103*, 1904.
(6) Miller, M. A.; Doye, P. L. J.; Wales, D. J. *Phys. Rev. E* **1999**, *60*, 3701.
(7) Berry, R. S.; Elmacı, N.; Rose, J. P.; Vekhter. B. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9520.
(8) Mezey, P. G. *Potential Energy Hypersurfaces*; Elsevier: Amsterdam, 1987.
(9) Becker, O. M.; Karplus, M. *J. Chem. Phys.* **1997**, *106*, 1495.
(10) Levy, Y.; Becker, O. M. *J. Chem. Phys.* **2001**, *114*, 993.
(11) Krivov, S. V.; Chekmarev, S. F.; Karplus, M. *Phys. Rev. Lett.* **2002**, *88*, 038101.
(12) Miller, M. A.; Wales, D. J. *J. Chem. Phys.* **1997**, *107*, 8568.
(13) Calvo, F.; Galindez, J.; Gadea, F. X. *J. Phys. Chem. A* **2002**, *106*, 4145.
(14) Hinde, R. J.; Berry, R. S. *J. Chem. Phys.* **1993**, *99*, 2942.
(15) Yurtsever, E.; Elmacı, N. *Phys. Rev. A* **1997**, *55*, 538.
(16) Stillinger, F. H.; Weber, T. A. *Science* **1984**, *225*, 983.
(17) Golub, G. H.; van Loan, C. F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, 1989.
(18) Aquilanti, V.; Cavalli, S. *J. Chem. Phys.* **1986**, *85*, 1355.
(19) Aquilanti, V.; Beddoni, A.; Cavalli, S.; Lombardi, A.; Littlejohn, R. *Mol. Phys.* **2000**, *98*, 1763.
(20) Aquilanti, V.; Beddoni, A.; Lombardi, A.; Littlejohn, R. *Int. J. Quantum Chem.* **2002**, *89*, 277.
(21) Aquilanti, V.; Lombardi, A.; Yurtsever, E. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5040.
(22) Mezey, P. G. *Shape in Chemistry*; VCH: Weinheim, Germany, 1993.
(23) Elad, M.; Tal, A.; Ar, S. HP Internal Report, 2000.
(24) Kuhn, H. W. *Nav. Res. Log. Q.* **1955**, *2*, 83.
(25) Hahn, P.; Grant, T.; Hall, N. *Eur. J. Oper. Res.* **1998**, *108*, 629.
(26) Thompson, K. C.; Jordan, M. T.; Collins, M. A. *J. Chem. Phys.* **1998**, *108*, 564.
(27) Rhee, Y. M. *J. Chem. Phys.* **2000**, *113*, 6021.
(28) Curotto, E.; Matro, A.; Freeman, D. L.; Doll, J. D. *J. Chem. Phys.* **1998**, *108*, 729.