

# Genetic Algorithms for Structural Cluster Optimization

Matthew D. Wolf\* and Uzi Landman

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332-0430

Received: March 11, 1998; In Final Form: May 1, 1998

Certain aspects of the methodology of genetic algorithms for global structural optimization of clusters were studied. Through systematic investigations of Lennard-Jones clusters with up to 100 atoms, several modifications were made to the genetic algorithm introduced by Deaven and Ho [*Phys. Rev. Lett.* **1995**, *75*, 288]. These modifications result in improved efficiency of the search procedure and in certain cases lead to determination of globally optimal structures that previous genetic algorithm studies have not found. The modifications include the following: twinning mutations, add-and-etch processes where a cluster of a given number of atoms is grown beyond that size and subsequently etched, and seeding of the initial parental population with selected structural motifs, in conjunction with randomly chosen configurations.

## I. Introduction

The development of optimization strategies (global optimization in particular<sup>1–3</sup>) is a subject of great importance in diverse fields, including the physical sciences (for example, protein folding and atomistic structural determination of (macro)-molecules and clusters), engineering (for example, design of VLSI circuits, industrial optimization, and control of regulation systems), and economics (for example, warehouse maintenance and optimal transportation and distribution routes). Generically, in optimization problems the overall performance of the system is represented by a multivariate function called the objective function; for example, the energy of a polyatomic aggregate with the atomic coordinates as variables. Optimal conditions are achieved when the objective function reaches its global extremum, i.e., minimum energy in the above example. However, for systems characterized by a large number of parameters, finding the extrema and, in particular, the global extremum is a vexing problem. The main difficulty is that the global extremum of a real multivariate function is actually a local property, thus requiring an exhaustive search. Furthermore, proving that the global extremum has indeed been found seems to be a rather unattainable task for most systems of interest.<sup>3</sup>

Several methodologies aimed at global optimization have been developed. These include gradient and Hessian methods,<sup>4</sup> stochastic and simulated annealing,<sup>5–7</sup> deterministic techniques,<sup>3</sup> genetic algorithms,<sup>8,9</sup> and other approaches (such as a basin-hopping method applied recently in a study of Lennard-Jones (LJ) clusters,<sup>10,11</sup> which is very similar to a Monte Carlo minimization method<sup>12</sup>).

As has been discussed in several recent papers,<sup>13–21</sup> genetic algorithms (GAs) offer a powerful tool for atomic and molecular cluster structural optimization. It is primarily due to the recent innovations in mating procedures<sup>13</sup> that the GA, long an important method in discrete optimization tasks,<sup>8,9</sup> became relevant to continuous variable optimization as well. Using such extended genetic algorithms, studies of the minimum-energy structures for C<sub>60</sub>,<sup>13</sup> the Thomson problem,<sup>22</sup> and Lennard-Jones (LJ) clusters<sup>14</sup> have been rather successful. Choosing LJ clusters as our model system, it is the goal of this study to examine in some detail the use of GAs for structural optimization and to

suggest several further extensions for improving the performance of GAs in such investigations.

In the discrete GA, a binary coding for the solution space is created first.<sup>8</sup> With this code, random strings are generated to serve as the initial generation. Each of these solutions is then tested (weighted) according to some (objective) function which reflects their “fitness” as solutions. The fittest solutions are then chosen as parents for the next generation. The parents are crossed (mated) in an analogue of sexual reproduction, yielding a new generation that should share some of the desirable traits of the parents. The process then repeats. This direct mimicry of natural selection has proven to be very efficient in many discrete optimization problems at finding, if not the absolute (global) minimum, then a sufficiently low local minimum for the application.

An innovative GA approach has been proposed<sup>13</sup> by Deaven and Ho (DH-GA), for structural (energetic) optimization of cluster systems. Rather than trying to create a complicated coding scheme that would allow one to use the traditional mating scheme of cutting and pasting binary strings, they suggested using the direct three-dimensional coordinate space representation of the clusters’ structures in conjunction with a particular mating procedure. In section II, we review the DH-GA algorithm, discuss several modifications to it, and present our results for LJ clusters of up to 100 atoms, which are in complete agreement with the results obtained in ref 10 (see in particular Table 1 in that reference), including certain cases that have not been found previously with the use of genetic algorithms. We conclude in section III.

## II. Genetic Algorithm and Application to Lennard-Jones Clusters

**II.A. The Genetic Algorithm and Modifications.** In the original DH-GA for a system of N atoms, one starts from a population of candidate structures whose configurations are randomly chosen subject to the constraints that no atoms lie on top of each other and that the chosen configuration is contiguous; that is, groups of atoms are close enough to interact. Following relaxation of these candidates to the nearest local minimum, a fraction of them are chosen as “parent” clusters with their relaxed energies serving as the fitness criteria. (The number

of parents,  $P$ , is taken to be small, though typically no smaller than  $P = 4$ .) These clusters are “mated” to form a new generation of clusters by choosing (in step i) a random plane passing through the center of mass of each of the parent clusters and cutting the parent clusters along those planes. In the following we call the two fragments thus generated for each cluster the “top” and “bottom” fragments, respectively. The planes are shifted for each parent so as to create two fragments with  $N/2$  atoms when  $N$  is even and two fragments with  $(N - 1)/2$  and  $(N + 1)/2$  atoms when  $N$  is odd. This step generates a total of  $2P$  fragments.

In step ii, all possible combinations of assembling new “offspring” clusters (each containing  $N$  atoms) from the  $P$  “top” fragments and the  $P$  “bottom” fragments are formed. For an even number of atoms, each fragment could mate with the  $2P - 1$  other fragments, since each of them contains  $N/2$  atoms. This is not the case for odd-numbered clusters. To assess the merits of our modifications to the GA on equal footing for both even and odd  $N$ , we only mate “top” with “bottom” for both parities of  $N$ .

Each child produced in this mating step is then relaxed to the nearest local minimum through a conjugate-gradient minimization, using the total energy as an objective function. Subsequently, a new generation is produced by selecting the  $P$  lowest energy clusters from the  $P^2$  combinations ( $P$  parents with  $P(P - 1)$  offspring). To preserve the population’s diversity, a minimum-energy difference  $\delta E$  between the members of the population is enforced. (That is, if two members of the population were to have energies within  $\delta E$  of each other, only one of them would be kept.)

Finally, in step iii, occasionally (at a rate of a few percent of the mating operation rate) mutated candidates are admitted into the population, regardless of their energy. Two mutation schemes have been suggested.<sup>13,14</sup>

We implemented first the original DH-GA<sup>13,14</sup> using the 6–12 LJ interatomic pair potential

$$v(r_{ij}) = 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right] \quad (1)$$

where  $r_{ij}$ ,  $2^{1/6}\sigma$ , and  $\epsilon$  are the distance between atoms  $i$  and  $j$ , the equilibrium LJ interatomic distance, and the equilibrium LJ well depth, respectively. In the following we use reduced units, where  $\sigma = \epsilon = 1$ . The energy of the  $i$ th atom in the cluster containing  $N$  atoms is given by  $E_i = 1/2 \sum_{j \neq i}^N v(r_{ij})$ , and the total energy of the cluster, given by  $E = \sum_{i=1}^N E_i$ , serves as an objective function. In the calculations reported here the population size was  $P = 4$  (although calculations were also done for  $P = 5$  and  $P = 6$ ), and the population diversity parameter was  $\delta E = 1 \times 10^{-2}$  (in reduced energy units).

Starting with this algorithm, we have experimented with several modifications aimed at improving the efficiency of the algorithm. We may classify these modifications as pertaining to the following: (i) mutations, (ii) initialization, and (iii) acceleration methods. We discuss each of these separately.

(i) *Mutations.* Most researchers working with the genetic algorithm tie the mutation frequency to a random probability, frequently around several percent of the mating operation rate. Instead, we choose to adjust the frequency of mutation according to the properties of the parental energy distribution. Whenever the population stagnates, we mutate one of the parents and place the mutated version in the parent list. We use two measures of stagnation. The first is triggered when the standard deviation of the parental energy distribution drops below a threshold value,

chosen to be 0.2 reduced energy units. The second criterion is activated when the parental energy distribution remains completely unchanged for three successive generations. The broadening of the search area in solution space (through mutations) is thereby tied to the trapping of the parents into “static” configurations. While in all cases mutations have been found to be triggered by the above criteria, it is advisable to include the option of a minimum mutation frequency of a few percent of the mating operation rate in cases where the stagnation criteria are not met.

The most basic mutation to perform on an atomic cluster is to select some random number of atoms and to take them on a random walk in configuration space,<sup>13</sup> while ensuring that they are not walked unphysically close to any other atom at each step. Despite its attractive simplicity, this method does not seem to dramatically alter the efficiency of the GA.

The next mutation that we implemented is a twinning mutation, motivated by the prevalence of twinned structures in clusters. A plane is randomly chosen through the center of mass of the cluster. One side of the cluster is then rotated by a random angle, and the two sides are then reassembled to form a mutated (twinned) cluster. Such mutations are rather efficient at moving the parent into a different area of solution space, as measured by changes in system characteristics such as number of nearest neighbors and average per-particle energy, and they improve the convergence time over the random walk mutation alone.

As a further class of mutations, we chose one that mimics certain physical processes operative in the formation of clusters and nanocrystallites. In particular, we developed etching mutations, implemented as follows.

First, a set number of particles  $M$  (typically  $M = 10$ ) are randomly placed in a shell surrounding the  $N$ -atom cluster. Precaution is taken to ensure that no pair of atoms are placed unphysically close to one another. The configuration of the cluster is then optimized (roughly) using a conjugate gradient minimization of the energy to generate a reasonably compact cluster. Through this procedure we achieve a workable approximation to an epitaxial deposition-and-growth process, which is expected to fill in any defects that may exist in the outer atomic shell structure of the parent cluster.

However, the epitaxial growth step leaves the mutated parent cluster with more atoms ( $N + M$ ) than the remainder of the parental clusters ( $N$ ). Consequently, the “growth” step is followed by “etching”. To this aim the per-atom energy for each atom in the cluster is calculated, and the highest energy particle is etched away. As a particle is removed, the energy of the cluster is reoptimized, and the etching process is repeated until the cluster returns to the original number of atoms ( $N$ ). Recognizing that these two steps can go in either order, we also created a mutation where the cluster is first etched and then experiences epitaxial growth. Additionally, the whole etching process can occur in sequence as described above, or it can be stretched over several generations of the GA if all parents are “grown” to the same size (i.e., the etching process is intertwined with the GA mating step).

(ii) *Initialization: Seeding.* To initialize the GA process, one may start from a parental population with random configurations (as in the DH-GA). While such choice is clearly without bias, it may, and in most cases does, lead to an increased computational effort. Our experience shows that even after an initial conjugate-gradient minimization ( $\sim 150$  steps) of randomly chosen initial configurations, the resulting initial population often consisted of some noncompact clusters.

Motivated by the above considerations, and in analogy with the physical processes of homogeneous nucleation and growth of materials aggregates, we have experimented with seeding procedures for initialization of the GA search. Such seeding is guided by the various structural motifs that are frequently encountered in studies of small- and medium-sized atomic aggregates.<sup>23–27</sup> For example, for clusters in the size range  $20 < N \leq 60$  and for  $P = 4$ , the initial parental population is seeded in the following manner: in addition to an  $N$ -atom randomly chosen parent, the population includes (three)  $N$ -atom clusters seeded with a 13-atom icosahedron, a 13-atom cuboctahedron, and a 13-atom decahedron, respectively. When seeding is used, we start with a population of  $P$  relaxed parental clusters; we found that starting from a larger initial population from which  $P$  clusters are selected, after relaxation, as described at the beginning of section II.A, has no discernible effect on the convergence to the global minimum when compared to the procedure outlined here. In the seeded clusters the remaining atoms ( $N - 13$ ) are placed at random within some region bordering the surface of the seed. When the number of parents  $P$  is greater than 4, this pattern (random, icosahedron, cuboctahedron, decahedron) is cyclically repeated. For other size ranges, the size of the seed is increased; for example, in the range  $60 < N \leq 147$  the seeds consist of a 55-atom icosahedron, a cuboctahedron, and a decahedron.

Conjugate-gradient minimization of the seeded clusters often results in more compact structures than minimization of the random clusters. We remark that while seeding may be regarded as introducing some bias into the GA, we rely on the presence of the randomly chosen configuration in the initial parental population, and on the action of the mutation operators described above, to alleviate undue bias in the GA search space.

In practice, we find it advantageous to grow the parental clusters (both random configuration and seeded ones) to a size of 10–20% larger than  $N$  and subsequently to etch the extra atoms as described above. This leaves a set of ( $P$ ) parental  $N$ -atom clusters, ready for the first mating procedure. Alternatively, one may begin the GA mating process before etching and perform the etching process gradually. We experimented with several etching schemes, all of them with essentially linear rates but having different slopes. Our best results (i.e., fastest convergence to the optimized structure) are for a scheme where the first atom is etched after three GA generations, and then in each subsequent generation an atom is removed (etched) until a cluster of  $N$  atoms is reached.

(iii) *Acceleration.* As aforementioned, after a mating step the heterozygous offspring are structurally optimized using a conjugate-gradient (CG) minimization of the energy. This is a time-consuming step, since it involves a large number of energy evaluations for each of the offspring. It has been suggested<sup>21</sup> that including information on the gradient of the total energy in the GA weighting scheme (objective function) might help to improve the convergence of the algorithm; in ref 21, a weighting function biased toward those distributions with a flat gradient was used. We have elected to utilize the gradient information from the CG minimization to perform the total minimization in two stages.

First, we perform an initial (approximate) minimization of all the offspring, but with a preset maximum number of CG evaluations (say 100). At this stage further optimization is performed only for the  $P$  heterozygous offspring selected as follows: (i) the norm of the gradient of each of the offspring generated by the mating process is subtracted from the roughly minimized energy of that structure; (ii) the  $P$  offspring clusters

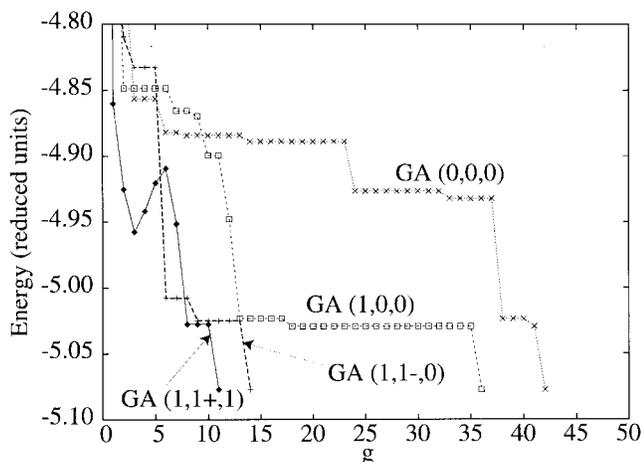
that have the lowest combined weight are selected for further minimization. The “best”  $P = 4$  clusters from both the parents and the children were then chosen based on their energies alone to become the next generation’s parents. Thus, gradient information was not included directly in the weighting function, but it was used as a time-saving measure in the process of deciding which cluster configurations were worthy of further minimization and study.

**II.B. Results: Lennard-Jones Clusters  $A_N$ ,  $N < 100$ .** (*i*) *Methodology.* The above modified GA was used by us for structural optimization of LJ clusters,  $A_N$  ( $10 \leq N \leq 100$ ). Prior to presenting these results, we illustrate the various modifications to the GA algorithm discussed above using  $A_{55}$  as an example. For the purpose of this illustration we add the various modifications and refinements progressively, allowing us to discern the role each plays in affecting the overall performance of the search algorithm. The code that we use in order to distinguish among the variants of the GA procedure is specified by the parameters  $(\alpha, \beta, \gamma)$  with the parameters taking the values 0 or 1 depending on the particular algorithm being used. The original DH-GA<sup>13,14</sup> described at the beginning of section II.A is assigned an index (0,0,0), and it includes initial parents ( $P = 4$ ) chosen from a population of relaxed random configurations, a mating process as described above in conjunction with CG minimization, a population diversity parameter of  $\delta E = 1 \times 10^{-2}$ , and two types of mutations (a random walk and a twinning mutation) which are triggered by stagnation criteria (described under “Mutations” in section II.A). This algorithm is used as the base of our notations.

When an “add-and-etch” mutation is included in the pool of possible mutations, we set  $\alpha = 1$ . When the initial population is completely composed of random parents,  $\beta = 0$ , while  $\beta \neq 0$  if seeding of the initial population is used. When seeding is used, it always comes in conjunction with an add-and-etch process. The seeding and accompanying add-and-etch processes are designated as follows:  $\beta = 1+$  or  $1-$ , indicating whether the etching is performed over several GA generations (+) or in one step before the first mating procedure (–). Finally, when gradients are included in the selection criterion  $\gamma = 1$ , while a purely energy-based criterion is represented by  $\gamma = 0$ . Thus, for example, a GA labeled by (1,1+,1) signifies an algorithm where “add-and-etch” mutations are used (in addition to random walk and twinning ones), the initial parental population is seeded in conjunction with a gradual add-and-etch process, and the selection criterion is based on both the energy and its gradient.

Typical results of energy per particle plotted versus the number of GA generations are shown in Figure 1 for  $A_{55}$ . These results were obtained using the original DH-GA (0,0,0) and several variants. It is evident that adding the “add-and-etch” mutation (1,0,0) improves the rate of convergence to the optimal structure (icosahedron) and that further significant enhancement of the convergence rate is achieved through seeding of the initial parental distribution in conjunction with an “add-and-etch” process (with the etch performed prior to the start of the GA mating procedure (denoted as algorithm (1,1–,0)). Finally, an even faster rate of convergence is achieved by performing the etch of the overgrown seeded initial population in a gradual manner, as well as including gradients in the selection criterion (i.e., the (1,1+,1) GA).

To illustrate the effect of the various improvements in a more statistical manner, we show in Figure 2a–d results for the variations in the number of GA generations needed to achieve the optimal structure for LJ clusters in the size range of  $40 \leq N \leq 50$ . The particular algorithms shown are the DH-GA



**Figure 1.** Comparison between the performance of the original DH-GA (denoted as GA (0,0,0)) and several variants (see text for the coding of the GA variants) for a LJ cluster with 55 atoms. In each case we display the per-particle energy (in reduced units) plotted vs the number of genetic algorithm generations,  $g$ . The number of generations required for convergence to the global optimal (icosahedral) structure corresponds to the last point given. Note the marked effect on the convergence rate of seeding in conjunction with add-and-etch processes (GA (1,1-,0) and GA (1,1+,1)).

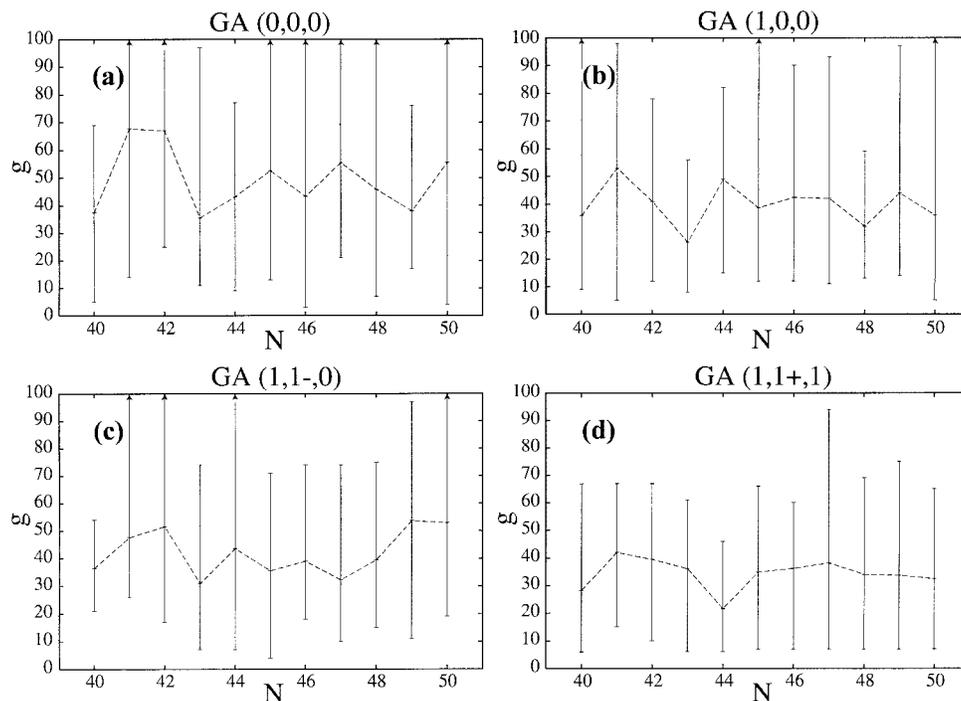
(0,0,0) algorithm (Figure 2a) and the variants (1,0,0), (1,1-,0), and (1,1+,1) (Figure 2, b, c, and d, respectively). In each case, the results that we show were obtained from 10 runs, and the vertical bars connect the smallest and largest number of GA generations that were required to achieve the optimal structure for each cluster size. When the latter exceeds 100 generations, the vertical line terminates with an arrow. The dashed line connects the (approximate) mean number of GA generations needed for optimization. Due to computational constraints, any

run that had not converged within 100 generations was terminated, and thus the mean appears artificially low for any cluster size that has such terminated (or “escaped”) runs.

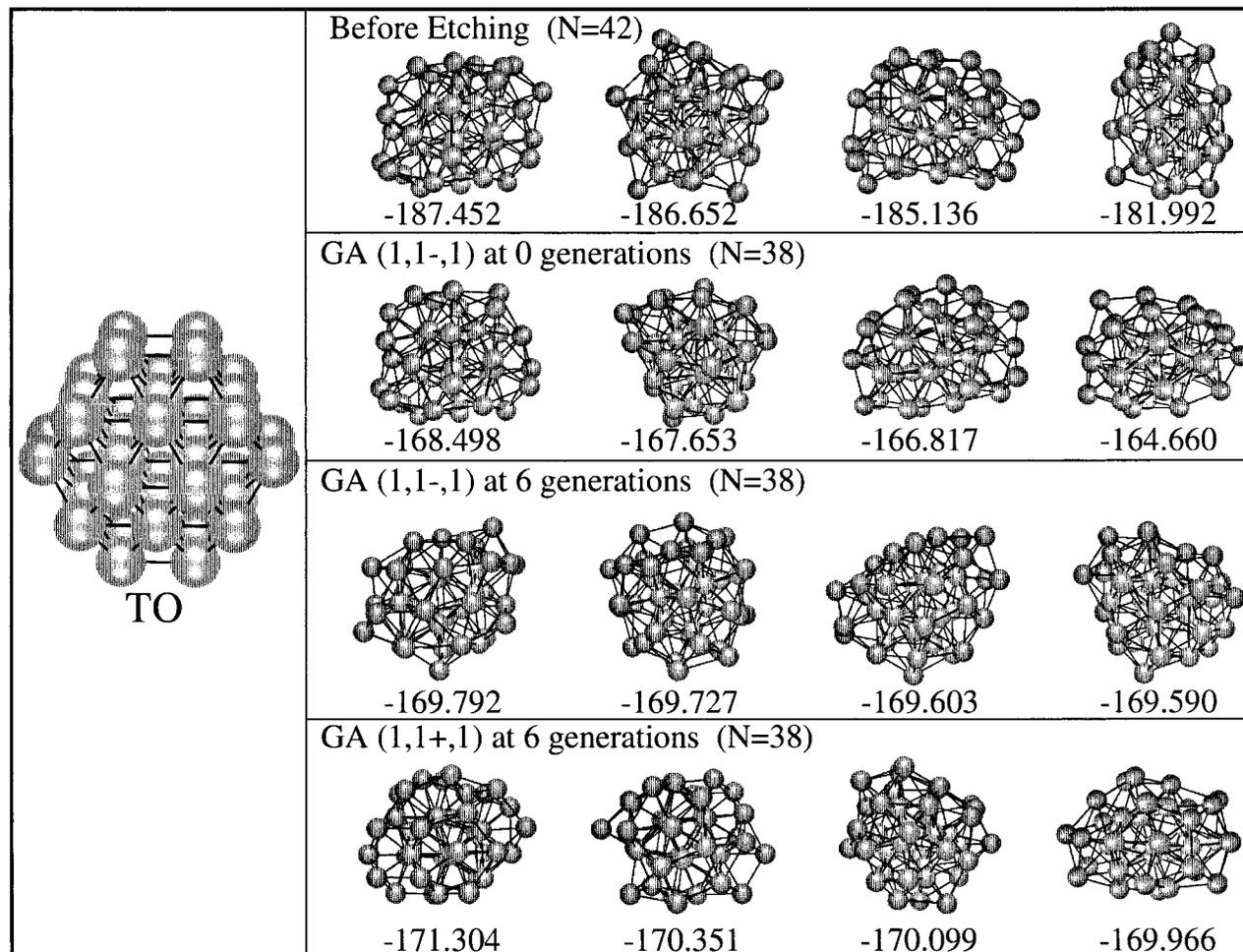
The merit of the various variants to the (0,0,0) GA can be assessed from the decreasing pattern in the average number of GA generations needed for convergence to the optimal structures. Furthermore, we observe that, for the majority of cluster sizes considered here, the (0,0,0) GA (i.e., the original DH-GA) experienced at least one run out of 10 which escaped (number of generations was greater than 100), while such cases are circumvented upon successive improvements of the algorithm (in particular see Figure 2d corresponding to the GA (1,1+,1) algorithm).

We conclude our discussion of the GA methodology developed here with an illustration of computations for  $N = 38$  pertaining to the seeding of the initial parental population (performed in conjunction with add-and-etch processes, as described above). As has been noted before, the energetically optimal structure (the fcc truncated octahedron (TO), shown on the left of Figure 3) is a difficult minimum to find through any automated global optimization method;<sup>10,11</sup> however, the optimal TO structure has been found for  $\text{Au}_{38}$ , using embedded atoms potentials via an “educated manual” exhaustive search (see Fig. 3 in ref 25 and Fig. 1 in ref 27). Consequently, illustration of the evolution of the parental clusters using our proposed GA variants, in particular (1,1-,1) and (1,1+,1), would serve to demonstrate certain aspects pertaining to these methods.

The initial seeded population ( $P = 4$ ) for both of these GA variants is grown to 42 atoms as shown in the top panel of Figure 3. In the second and third panels from the top, we show the parental population for the (1,1-,1) GA immediately after the fast etch (prior to GA mating steps) and the parental population after six subsequent GA generation, respectively. In



**Figure 2.** Information pertaining to the number of genetic algorithm generations required for convergence to the global optimal structures for LJ clusters,  $A_N$ , with  $40 \leq N \leq 50$ . For each cluster size the results were obtained from 10 independent runs. The dashed lines connect the mean number of generations,  $g$ , required for convergence, and the upper and lower values of each of the vertical bars correspond to the largest and smallest number of generations (out of 10 runs) for which convergence was achieved. An arrow at the top of a vertical bar indicates failure to achieve convergence after 100 generations. The original DH-GA (labeled GA (0,0,0)) and the GA variants are marked at the top of each panel (see text for the coding scheme of the GA variants).



**Figure 3.** Illustration of the effect of seeding in conjunction with add-and-etch operations on the initial stages of a genetic algorithm search for  $A_{38}$  (the global optimal, truncated octahedral (TO), structure of the cluster is shown on the left). A seeded population ( $N = 42$ ) of initial parents ( $P = 4$ ) is shown in the top panel; going from left to right, the parents were seeded with a 13-atom icosahedron, a 13-atom cuboctahedron, and a 13-atom decahedron, and one (on the right) was not seeded (i.e., the positions of all atoms were randomly chosen). The configurations at the top panel were recorded after the conjugate-gradient minimization as described in the text. The configurations shown in the second panel from the top correspond to (fast) etching of the initial configurations to 38 atoms before starting the GA mating process (i.e., following the GA (1,1-,1) variant). In the third panel from the top we display the population after six genetic algorithm generations (starting from the parents shown in the second panel). In the bottom panel, we show the population after six genetic algorithm generations, obtained from the initial (top panel) parents through a (slow) etching process (i.e., the GA (1,1+,1) variant). The total energies (in reduced units) are given for each of the cluster configurations. (Note that the number of atoms in the top panel is  $N = 42$ , while  $N = 38$  in all the other panels.)

the fourth (bottom) panel we show the parental population after six GA generations for the (1,1+,1) GA. (Since the (1,1+,1) GA uses a gradual etch interwoven with the GA mating operations, with the first etch process performed after three initial GA generations and one atom etched after each subsequent generation, it takes six GA generations to etch the clusters from  $N = 42$  to  $N = 38$ .) The total energies of the pictured clusters are included in the figure. We note that with the (1,1+,1) GA the configurations (after six generations) are of lower energy than those generated via the (1,1-,1) GA; while this example illustrates a commonly observed trend, it should not be inferred that it occurs at all instances.

Note that although the clusters resulting from the (1,1+,1) GA are rather compact, none of them shares any strong structural resemblance to the absolute (global) minimum for  $N = 38$ . While eventually the GA converges to the optimal (TO) structure, a study of the structures in the energetic neighborhood of the TO minimum should help us to clarify some of the inherent difficulties encountered in many instances in structural and energetic optimization of atomic and molecular clusters. Such cases consist of systems for which, in certain regions of

the multidimensional potential energy surface (PES), the total energy of the cluster only weakly differentiates between isomeric structures, even when these belong to different structural motifs. The energy gap between the TO and its nearest *structurally related* isomer, which corresponds to a single vacancy on the surface of the TO cluster with an adsorbed atom at a 4-fold site (the {100} facet directly opposite the vacancy), is quite large, i.e., about 2.5 reduced energy units. However, within that energy gap, we have documented at least 14 other local minima that display multitwinned fcc crystallites of the sort seen in icosahedral and decahedral structures, rather than the pure fcc crystal which typifies the TO. Thus, with great likelihood a GA, or any other global minimization search, will spend a substantial amount of time searching in the neighborhood of these low-lying, structurally dissimilar isomers.

For the GA, this is particularly troublesome, since the GA's basic method involves recombining (mating) the fragments of the preceding generation to make the following one. If there is no way in which the structures of one set of parents can be reassembled into an energetically competitive isomer that is structurally related to the global minimum, then the algorithm

must depend on the mutation operator to modify some parental cluster for convergence to the global minimum to occur.

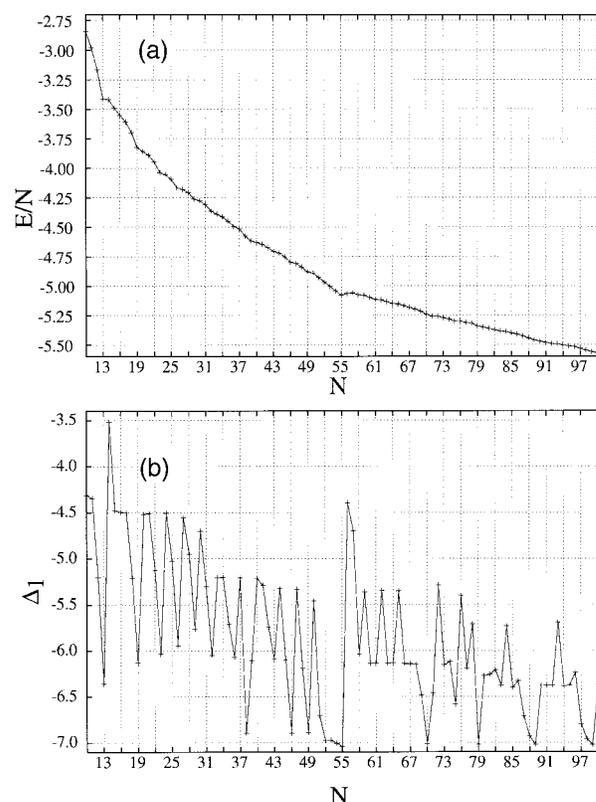
Alternatively, the parental population may be increased to the point where the number of parents,  $P$ , is greater than the number of intervening local minima. However, there is no a priori way to know how many parents this would require. Furthermore, the number of clusters to be minimized at each generation goes up as  $P(P - 1)$ , so simply increasing  $P$  to arbitrarily large values is computationally inefficient.

Given that many such cluster problems fall into the class of computationally hard problems, it becomes important to explore the added information that can be extracted from GA studies, pertaining to the nature of the potential energy surface (PES) of a system, even if it cannot be ensured that the GA has found the global minimum. In this respect, the genetic algorithm has a distinct advantage over many alternative optimization methods. Through the maintenance of parental and child energy lists, the GA can very efficiently explore the low-lying regions of the PES. Since each child is guaranteed to share at least some of the geometrical characteristics of the parental local minima, relatively little time will be spent exploring the uncompetitive regions of the PES. For studies of ground-state behavior at finite temperature, such local minima lists can be valuable. Furthermore, by examining the structure of the clusters within the energy lists, new structural classes may be found. It was through just such an analysis that we discovered what we believe to be a new closed-shell structure for the 55-atom cluster (see Appendix).

(ii) *Optimal Structures of LJ Clusters;  $N \leq 100$ .* Clusters of atoms interacting via LJ pair potentials have been intensively investigated theoretically<sup>10,11,13,23,28–39</sup> and experimentally, and they currently serve as a generic test system for global optimization algorithms. This system, which is characterized by a simple (nonconvex) objective function (i.e., the total energy expressed as a sum of all the pair interactions in the cluster), is rather challenging for global optimization since it possesses a very large number of geometrically distinct isomeric structures, growing with  $N$  as<sup>29</sup>  $m(N) \sim \exp(0.3572N + 0.0286N^2)$ , predicting  $m(N)$  on the order of  $10^{140}$  for a cluster with  $N = 100$  atoms. Such inflation in the number of geometrically distinct isomers, and consequently the number of local minima on the multidimensional potential energy surface, and the fact that LJ interactions are often used to describe nonbonded interactions in protein molecules lead one to suggest that there may be certain qualitative similarities in the potential energy landscape of systems characterized by these potentials and that investigations and methodologies developed for global optimization of atomic LJ clusters may be of some relevance to protein folding studies.

Most early theoretical structural studies of LJ clusters were based on the idea of growth sequences,<sup>28</sup> that is, sequential buildup of structural candidates of size  $N$  by adding an atom to a previously optimized structure of size  $N - 1$ , followed by relaxation to the nearest local minimum. Hessian matrix analysis was then used to determine the stability of alternative cluster structures. Using such methodology, early explorations<sup>28</sup> were based on a polytetrahedral growth sequence up to  $N = 66$ , with the starting seed being a regular tetrahedron ( $N = 4$ ).

Icosahedral packing sequences, constructed by adding successive icosahedral shells around a central atom, were introduced in 1962 by Mackay<sup>40</sup> on the basis of geometrical arguments without the use of a potential function. This idea provided the basis for later work<sup>31,32</sup> where icosahedral structures were investigated using the LJ interaction, resulting in a list<sup>32</sup> of



**Figure 4.** (a) Per-particle energies  $E/N$  (in reduced units) and (b) first energy difference,  $\Delta_1 = E(N) - E(N - 1)$ , for globally optimized LJ clusters,  $A_N$ , plotted vs  $N$  for  $10 \leq N \leq 100$ .

icosahedrally based structures in the size range  $12 \leq N \leq 147$ . Since then it has been found, using various methods, that while the majority of optimal structures in the above range are indeed of an icosahedral motif, there are also some that are made of face-centered-cubic lattices of particular morphologies (i.e., TO at 38) and some that follow the Marks' decahedral<sup>41</sup> (m-Dh) pattern (for a detailed discussion of Marks' decahedra see ref 23). In this context, it is worth remarking that such studies, including comparative investigations<sup>23</sup> of optimal cluster structures using continuum energetics (Wulff construction) and atomistic structural optimization, have been made for both LJ and metal clusters. Since written accounts of these studies are available,<sup>23</sup> we will focus our discussion in the following mainly on those using the GA as the global optimization method, which is the topic of this paper.

The energies (per atom) and first energy differences  $\Delta_1 = E(N) - E(N - 1)$  of the structurally optimal LJ clusters  $A_N$  for  $10 \leq N \leq 100$ , found by us via application of the GA (1,1+,1), are shown in Figure 4. All these energies and the corresponding structures are identical to those listed in Table 1 of ref 10. From Figure 4 it is evident that while the total energy of the LJ clusters varies with the number of atoms in a rather monotonic manner, it exhibits mild discontinuities at certain sizes; in particular, (i) at  $N = 13$  and  $55$  corresponding to filled icosahedral shells (note the flat region for  $52 \leq N \leq 55$  in Figure 4b, indicating a "basin of attraction" in the multidimensional configuration space culminating at  $N = 55$ ) and (ii) at or near  $N = 19, 38, 70, 75, 79, 89$ , and  $98$ .

In a recent study<sup>14</sup> using the DH-GA (denoted by us above as GA (0,0,0)) several optimal configurations with lower energy than published before were found for LJ clusters at  $N = 38, 65, 69, 76, 88$ , and  $98$ . These configurations were also found by us<sup>42</sup> using the GA (1,1+,1); in addition, optimal structures

of slightly lower energy than those found previously using genetic algorithm optimization were determined by us for  $N = 69$  and  $N = 75-78$ . Most importantly, the optimal configurations found by us using the GA (1,1+,1) were generated with a significant savings in computation time; for example, using our GA (1,1+,1) convergence to the TO optimal structure of  $N = 38$  was typically achieved in 25–50 genetic algorithm mating steps, which is a 3–4-fold faster convergence rate than that quoted for this cluster size in ref 14.

### III. Summary

In this study, we have focused on the methodology of genetic algorithms for global structural optimization of clusters and illustrated our modifications through an investigation of LJ clusters,  $A_N$ , for  $N \leq 100$ . Starting from the Deaven and Ho algorithm<sup>13</sup> (DH-GA), we have explored various variants aimed at improving the efficiency of the GA search. These new variants include the following: new mutations (twinning and add-and-etch), seeding of the initial parental population, and certain (gradient-based) acceleration procedures. Through systematic investigations we have assessed and demonstrated the merits of these GA variants and the higher efficiency which they bring in reducing significantly the number of GA generations to achieve convergence to the globally optimized structures.

Through application of our new GA procedure to LJ clusters, we determined the globally optimal structures for clusters with up to 100 atoms, obtaining results in agreement with those determined via other search and optimization methods (see in particular Table 1 in ref 10), including certain cases<sup>42</sup> that were not found in previous GA studies (see e.g. ref 14).

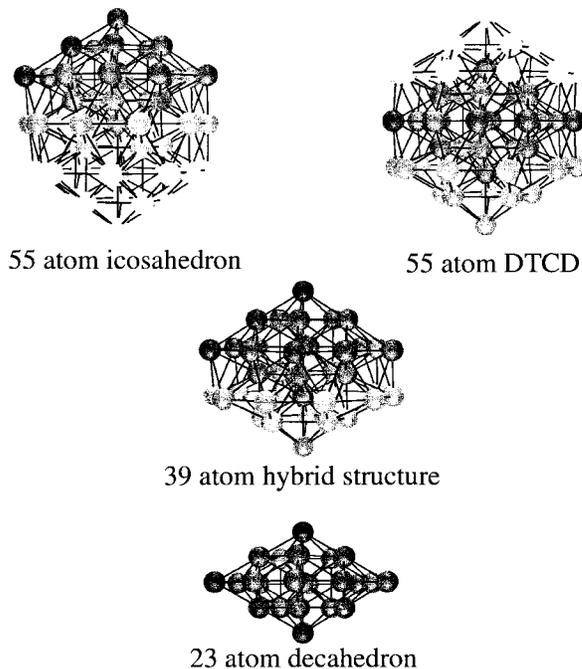
**Acknowledgment.** Useful discussions with Dr. C. L. Cleaveland are gratefully acknowledged. This work is supported by the U.S. Department of Energy Grant No. FG05-86ER45234. Calculations were performed at the Georgia Institute of Technology Center for Computational Materials Science.

### Appendix: The DTCD

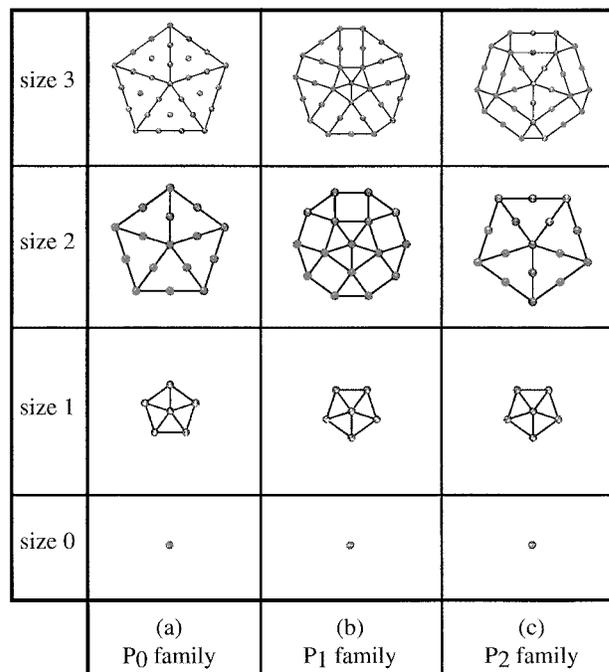
During our investigations of  $A_{55}$ , we noticed that convergence to the global minimum structure (Mackay icosahedron) was often somewhat hindered by a low-lying structurally distinctive isomer, which we have named the double twin-capped decahedron (DTCD). There are many local minima lying in the energy interval between the DTCD and the optimal 55-atom icosahedron which are structurally directly related to the icosahedron and which therefore provide a direct pathway to finding the global minimum (unlike the case of the 38-atom clusters discussed in section II.B(i)). Although overall convergence was therefore not significantly impacted, we investigated further the properties of the DTCD structure.

Like the icosahedron and the decahedron at  $N = 55$ , the DTCD can be viewed as being composed of multitwinned fcc crystallites. In fact, as seen in Figure 5, the icosahedron and the DTCD have many structural properties in common. We have found that the easiest way to explain the structural similarities is through a decomposition of the icosahedron by way of a series of vertex caps.

Each of the vertex caps which we use in our construction belongs to one of a set of infinite families, denoted here by  $P_0$ ,  $P_1$ ,  $P_2$ , and so on. Within each family, the caps come in sizes  $t$  from  $t = 0$  (a single atom) to  $t = \infty$ . The caps in family  $P_0$  we have called "decahedral caps" (Figure 6 (a)), since stacking such caps of successive sizes yields either a perfect or an Ino's

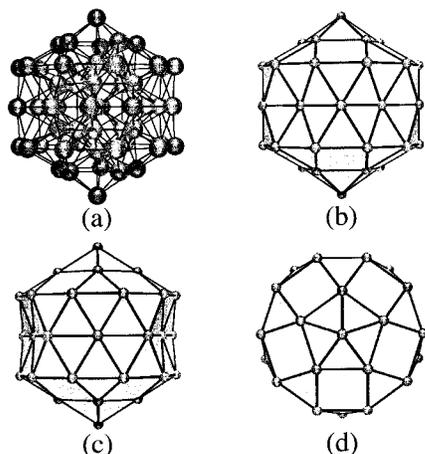


**Figure 5.** Seed for both the icosahedron and the DTCD is the 23-atom decahedron (the bottom dark gray structure). In the next growth step a complete twin cap ( $P_1$  cap of size 2 (see text and Figure 6 for the cap classification scheme)) is added to the bottom of the decahedron (the light gray atoms). In the final step for the icosahedron, a decahedral cap rotated by  $36^\circ$  (or a  $P_2$  cap of size 2) is added (bright white atoms) onto the bottom of the 39-atom structure. For the DTCD another twin cap ( $P_1$  cap of size 2) is added (bright white atoms) to the top of the 39-atom structure, yielding a plane of reflection symmetry through the center atom.



**Figure 6.** Projections onto the plane of the four smallest sizes of the first three families of pentaradial vertex caps. Note the obvious geometric extension to both larger sizes and higher order families.

decahedron (see description in ref 23, Figure 4); the generalization to Marks' decahedra is straightforward but not relevant to this discussion. The second family  $P_1$  we have called "twin caps" (Figure 6 (b)). The name, and the caps' structures, can be easily understood in the following context. A Mackay



**Figure 7.** The DTCD. View (a) shows the internal atomistic construction (discussed in Figure 5), while view (b) only shows the geometrical surface features. View (c) shows the geometrical surface rotated by  $36^\circ$  compared to (b), and (d) shows the structure in (b) viewed from above. Note that the geometrical surfaces shown do not correspond to a convex hull. The alternating reentrant faces along the circumference and the joining of the squares on the twin caps to the upper pentagonal pyramid are concave.

icosahedron can be viewed as a collection of 20 strained fcc crystals cut along their  $\{111\}$  planes. If we were to cap one of the vertices of this icosahedron with a  $P_1$  (twin) cap, then each of the triangular faces of the twin cap would sit on the five adjacent fcc crystallite faces as though the triangular facet of the icosahedron was a twinning plane in the fcc crystal.

The caps in the families  $P_n$  for  $n \geq 2$  can be easily generated from the following recursive description (refer to Figure 6 (c)). For a cap of size  $t$  in family  $P_n$ , the atoms in the lower triangular faces (if any) lie in fcc stacking order with the atoms of the lower triangular faces of the cap of size  $t$  in  $P_{n-1}$ . The rectangular regions (which correspond to a strained  $\{110\}$  direction in the fcc crystallites) extend between these triangles. Finally, the area from the central vertex to the long sides of the rectangles is filled with a  $P_0$  (decahedral) cap, rotated by  $36^\circ$ , of size  $n - 1$  if  $t > n - 1$  or of size  $t$  if  $t \leq n - 1$ . (If  $t \leq n - 1$ , then there will not be any lower triangular or rectangular pieces.) Since the 55-atom clusters (decahedral or icosahedral) only call for caps of size 2 or smaller, all of these higher order families ( $P_n$  for  $n \geq 2$ ) appear indistinguishable from a  $P_0$  (decahedral) cap rotated by  $36^\circ$ .

From Figure 5, we see that the icosahedron can be constructed by applying successive layers of both the  $P_0$  caps and the  $P_1$  caps. The root 23-atom decahedron can be viewed as a  $P_0$  (decahedral) cap of size 0 (a single atom), capped by a  $P_0$  cap of size 1, and then capped by a  $P_0$  cap of size 2. The addition of a  $P_1$  (twin) cap of size 2 brings the number of atoms to 39. Up to this point the constructions of the 55-atom Mackay icosahedron and the DTCD are identical. To create the icosahedron, a  $P_2$  cap of size 2 (or, equivalently, a decahedral cap of size 2 rotated by  $36^\circ$ ) is placed over the  $P_1$  (twin) cap of the 39-atom hybrid structure. However, by placing a second twin cap over the opposite side of the 39-atom hybrid structure, the DTCD is created.

Although its structure is nonconvex, the DTCD has some compelling geometrical and symmetrical properties (see Figure 7). Because the second  $P_1$  (twin) cap restores the mirror plane through the central atom, the DTCD belongs to the decahedral symmetry group. However, the distribution of strain within the crystallite is distinct from that of the Ino's decahedron at 55

atoms (denoted as  $m = n = 3$  in ref 23), as well as that of the 55-atom icosahedron.

While the DTCD is not energetically lower than the Mackay icosahedron, it is the closest closed-shell structure at 55. Both the cuboctahedron and the Ino's decahedron at 55 lie substantially higher in energy. This and the fact that the strain within the DTCD is distributed in a manner distinct from that in both the icosahedron and decahedron make the DTCD a promising candidate for potentials that are less tolerant of strain than the LJ potential. As previously found,<sup>37</sup> there is a transition from icosahedral structures to decahedral structures to strain-free structures (like fcc or bcc) as the range of the interatomic (pairwise) potentials become longer (and hence less tolerant of high strain). We therefore suggest that, in that sequence of transitions, there may be places where the DTCD or other related multitwinned structures become energetically favorable over their more well-known kin.

## References and Notes

- (1) *State of the Art in Global Optimization: Computer Methods and Applications*; Floudas, C. A., Pardalos, P. M., Eds.; Kluwer: Dordrecht, The Netherlands, 1996.
- (2) *Global Minimization of Nonconvex Energy Functions: Molecular Conformations and Protein Folding*; Vol. 23 of Series in Discrete Mathematics and Theoretical Computer Science (DIMACS); Pardalos, P. M., Shalloway, D., Xue, G., Eds.; American Mathematical Society: Providence, RI, 1996.
- (3) Barhen, J.; Protopoulos, V.; Reister, D. *Science* **1997**, *276*, 1094 and references therein.
- (4) Goto, H.; Osawa, E. *J. Mol. Struct. (THEOCHEM)* **1993**, *285*, 157.
- (5) Kirpatrick, S.; Gelatt, Jr., C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671.
- (6) van Laarhoven, P. J. M.; Aarts, E. H. L. *Simulated Annealing: Theory and Applications*; Reidel: Dordrecht, The Netherlands, 1987.
- (7) Aluffi-Pentini, F.; Parisi, V.; Zirilli, F. *J. Optimization Theory Appl.* **1985**, *47*, 1.
- (8) Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: Reading, MA, 1989.
- (9) *Genetic Algorithms and Simulated Annealing*; Davis, L., Ed.; Pitman: London, 1987.
- (10) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111.
- (11) Doye, J. P.; Wales, D. J. *Phys. Rev. Lett.* **1998**, *80*, 1357.
- (12) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611.
- (13) Deaven, D. M.; Ho, K. M. *Phys. Rev. Lett.* **1995**, *75*, 288.
- (14) Deaven, D. M.; Tit, N.; Morris, J. R.; Ho, K. M. *Chem. Phys. Lett.* **1996**, *256*, 195.
- (15) Mestres, J.; Scuseria, G. E. *J. Comput. Chem.* **1995**, *16*, 729.
- (16) Pullan, W. J. *Comput. Phys. Commun.* **1997**, *107*, 137.
- (17) Xiao, Y.; Williams, D. E. *Chem. Phys. Lett.* **1993**, *215*, 17.
- (18) Smith, R. W. *Comput. Phys. Commun.* **1992**, *71*, 134.
- (19) Brodmeier, T. E.; Pretsch, E. *J. Comput. Chem.* **1994**, *15*, 588.
- (20) McGarragh, D. B.; Judson, R. S. *J. Comput. Chem.* **1993**, *14*, 1385.
- (21) Gregurick, S. K.; Alexander, M. H.; Hartke, B. *J. Chem. Phys.* **1996**, *104*, 2684.
- (22) Morris, J. R.; Deaven, D. M.; Ho, K. M. *Phys. Rev. B* **1996**, *53*, R1740.
- (23) Cleveland, C. L.; Landman, U. *J. Chem. Phys.* **1991**, *94*, 7376 and references therein.
- (24) 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.; Landman, U. *Adv. Mater.* **1996**, *5*, 428.
- (25) Whetten, R. L.; Khoury, J. T.; Alvarez, M. M.; Murthy, S.; Vezmar, I.; Wang, Z. L.; Cleveland, C. L.; Luedtke, W. D.; Landman, U. *The Chemical Physics of Fullerenes 10 (and 5) Years Later*; Andreoni, W., Ed.; Kluwer: Dordrecht, The Netherlands, 1996; p 475.
- (26) Cleveland, C. L.; Landman, U.; Shafiqullin, M. N.; Stephens, P. W.; Whetten, R. L. *Z. Phys. D* **1997**, *40*, 503.
- (27) Cleveland, C. L.; Landman, U.; Schaaf, T. G.; Shafiqullin, M. N.; Stephens, P. W.; Whetten, R. L. *Phys. Rev. Lett.* **1997**, *79*, 1873.
- (28) Hoare, M. R.; Pal, P. *Adv. Chem. Phys.* **1971**, *20*, 161.
- (29) Hoare, M. R. *Adv. Chem. Phys.* **1979**, *40*, 49.
- (30) Hoare, M. R.; McInnes, J. A. *Adv. Phys.* **1983**, *32*, 791.
- (31) Farges, J.; de Feraudy, M. F.; Raoult, B.; Torchet, G. *J. Chem. Phys.* **1986**, *84*, 3491.
- (32) Northby, J. A. *J. Chem. Phys.* **1987**, *87*, 6166.

- (33) Raoult, B.; Farges, J.; de Feraudy, M. F.; Torchet, G. *Philos. Mag. B* **1989**, *60*, 881.
- (34) Maranas, C. D.; Floudas, C. A. *J. Chem. Phys.* **1992**, *97*, 7667.
- (35) Coleman, T.; Shalloway, D.; Wu, Z. *J. Global Optim.* **1994**, *4*, 171.
- (36) Xue, G. L. *J. Global Optim.* **1994**, *4*, 425.
- (37) Doye, J. P. K.; Wales, D. J.; Berry, R. S. *J. Chem. Phys.* **1995**, *103*, 4234.
- (38) Sloane, N. J. A.; Hardin, R. H.; Duff, T. D. S.; Conway, J. H. *Discrete Comput. Geom.* **1995**, *14*, 237.
- (39) Leary, R. H. *J. Global Optim.* **1997**, *11*, 35.
- (40) Mackay, A. L. *Acta Crystallogr.* **1962**, *15*, 916.
- (41) Marks, L. D. *Philos. Mag. A* **1984**, *49*, 81.
- (42) For  $N = 69$  and  $N = 78$ , the global optimal structures belong to the icosahedral motif. The structures for these clusters found by us (in agreement with those given in ref 10, Table 1) are lower in energy than those determined in ref 14; the structural differences between local minimum

structures given in ref 14 and the globally optimal ones, thought small (involving the displacement of a single atom), yield lower energies (i.e., we find total energies of  $-359.882\,566$  and  $-414.794\,401$  for  $N = 69$  and  $N = 78$ , respectively, while the energies given in ref 14 for these sizes are  $-354.845\,488$  and  $-414.681$ , in reduced energy units). The global minimum-energy configurations for clusters with  $N = 75-77$  belong to the truncated (Marks) decahedral motif, which has been discussed previously for gold<sup>26,27</sup> as well as LJ<sup>10</sup> clusters. The structures found for these clusters in ref 14 are (higher energy) isomeric forms belonging to an icosahedral motif. Furthermore, we have found (in agreement with ref 10, Table 2) that the results given in ref 14 for these local minima are not the optimal ones for these icosahedral isomers; total energies determined by us, as well as in ref 10, for the optimal icosahedral isomers of  $N = 75, 76$ , and  $77$  are  $-396.282\,243$ ,  $-402.384\,580$ , and  $-408.518\,265$ , compared to  $-396.037$ ,  $-402.368\,548$ , and  $-408.463$  given in ref 14. It is of interest to note that the less optimal isomers determined in ref 14 also occurred in our GA search along the path to the global optimal structures.