

An Efficient Deformation-Based Global Optimization Method (Self-Consistent Basin-to-Deformed-Basin Mapping (SCBDBM)). Application to Lennard-Jones Atomic Clusters

Jarosław Pillardy,[†] Adam Liwo,^{†,‡} and Harold A. Scheraga^{*,†}

Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, and Department of Chemistry, University of Gdańsk, ul. Sobieskiego 18, 80-952 Gdańsk, Poland

Received: August 3, 1999; In Final Form: September 17, 1999

A recently proposed method to surmount the multiple-minima problem in protein folding is applied here to global optimization of Lennard-Jones atomic clusters. The method, self-consistent basin-to-deformed-basin mapping (SCBDBM), locates a group of large basins containing low-energy minima (hereafter referred to as *superbasins*) in the original energy surface by coupling the superbasins in the original surface to basins in a highly deformed energy surface (which contains a significantly reduced number of minima, compared to the original rugged energy surface). Various kinds of deformation based on the distance scaling method (DSM) have been tested. The method was able to locate all the lowest-energy structures of Lennard-Jones atomic clusters with a size of up to 100 atoms, except for clusters of 75–77 atoms. In these cases, the method found the previously known second-to-the-lowest energy structures.

1. Introduction

The multiple-minima problem (the existence of many minima in the multidimensional surface of a multidimensional function) is encountered in essentially all areas of theoretical chemistry, physics, and engineering, as well as in many other branches of science and economics. In theoretical chemistry, the interest in finding the global minimum of a potential energy surface arises from the fact that it plays a very important role in locating the thermodynamically most stable structure. For example, according to the *thermodynamic hypothesis* formulated by Anfinsen,¹ the three-dimensional structure of a native protein corresponds to the global minimum of its free energy surface. This leads to theoretical predictions of structure using empirical potential functions whose parameters may be obtained from first principles or from experimental data. Because of its importance, much effort has been devoted to surmount the multiple-minima problem but, despite the great progress that has been achieved, it still remains unresolved. In particular, theoretical conformational analysis of macromolecules, investigations of the spatial structures of many-atom or many-molecule clusters, and theoretical prediction of crystal structures are at best very difficult to treat because of the huge number of local minima in the corresponding energy surfaces.^{2,3} It must be strongly emphasized that, for global optimization to succeed in predicting the properties of real systems, it is necessary to have an accurate potential energy function.

Since many different global optimization algorithms have already been proposed and applied to various systems, there is a need for a general and simple way to compare their effectiveness. The simplest way to carry out such a comparison is to apply any newly proposed (or improved) method to a standard system whose global optimization properties are well-

known. Lennard-Jones (LJ) atomic clusters represent such a system; since the mathematical form of the potential energy function is very simple (see eq 1), it is relatively easy to apply most global optimization methods. It is not easy, however, to locate the global minimum for a sufficiently large cluster. The number of local minima in LJ clusters grows exponentially. It has been estimated⁴ that a cluster of 55 LJ particles has at least 10^{10} geometrically different minima; this number grows to about 10^{60} for a cluster of 147 particles. LJ clusters have been thoroughly researched in the past years, and the most likely candidates for global minima are known^{5–7} (an up-to-date database of currently known lowest-energy structures of LJ clusters is accessible on the Internet⁸). The potential energy between atoms in an LJ cluster has the form

$$E = \sum_{ij} \epsilon \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_0}{r_{ij}} \right)^6 \right] \quad (1)$$

where ϵ is the pair equilibrium well depth and r_0 is the pair equilibrium distance (reduced units, i.e., $r_{0i} = 1$ and $\epsilon = 1$, are used in this work).

The majority of LJ_N clusters of size from 10 to 150 atoms have global-minimum structures based on the Mackay⁹ icosahedron, the complete icosahedron being formed for $N = 13$, 55, and 147. For intermediate sizes, the structure consists of a complete icosahedron core of the largest size possible for the given number of atoms, covered with a layer of atoms packed in order to minimize the potential energy. It is worth noting, however, that the pentagonal symmetry of the Mackay icosahedron is not possible in an infinite crystal; therefore, sooner or later, large LJ clusters will exhibit the hexagonal symmetry of the face-centered-cubic (fcc) crystal structure (which is observed experimentally). This change may be sharp (i.e., for clusters larger than a given size, all clusters will exhibit fcc symmetry), or may be gradual, so that the probability of the cluster showing fcc symmetry will grow from a very low value,

* To whom correspondence should be addressed. Phone: (607) 255-4034. Fax: (607) 254-4700. E-mail: has5@cornell.edu.

[†] Cornell University.

[‡] University of Gdańsk.

through 50%, and finally to 100%. For clusters of a size ranging from 10 to 150 atoms, only a few clusters have been found that have global minima that are not based on the Mackay icosahedron. The lowest-energy structure of the LJ₃₈ cluster is an fcc truncated octahedron;^{10,11} for $N = 75-77$, 102-104, structures have been found^{11,12} which are based on Marks' decahedron.¹³

Another possible structure for large Lennard-Jones clusters is the hexagonal-close-packing (hcp)^{14,15} one, which is very similar to fcc. The absence of the hcp form in rare-gas crystals is linked to the mechanism of crystal growth;¹⁵ however, it is quite possible to encounter them as low- or the lowest-energy structures for very large Lennard-Jones clusters using global optimization techniques.

In order to assess the potential properly for a given global optimization method applied to LJ clusters, it is necessary to know how the relative difficulty in finding the global minimum depends on the size of a cluster. In general, the difficulty grows significantly with the size of the cluster because of an increase in the number of local minima. A good procedure to estimate the relative difficulty is to carry out a large number of local minimizations starting from randomly chosen points in the configurational space; this procedure can indicate when the probability of randomly finding the global minimum is high because of a low number of minima or a large relative size of the global-minimum basin. The results of randomly started local minimizations reported in Table 1 clearly show that finding global minima of clusters smaller than 30 atoms is very easy and could be accomplished with only moderate numerical effort without using a sophisticated global optimization procedure. Local minima on the potential energy surface (PES), however, are organized hierarchically, forming large families of minima interconnected through relatively low barriers, but some families are often separated by very high energy barriers and long topological distances.¹⁶ When only one such family exists on the PES, the task of finding the global minimum energy is relatively easy, even for a large number of atoms in the cluster. A good example of this behavior is LJ₅₅, for which only one large family of nearby minima exists.¹⁷ On the other hand, clusters for which the PES exhibits several different families of structures are especially difficult for global optimization; examples are LJ_N for $N = 38$ and $75-77$.¹⁶ Taking the current computational capabilities of modern computers into account, it is clear that the assessment of any general global optimization method should be evaluated for LJ clusters having at least 35 atoms (preferably larger than 50 atoms); the clusters LJ_N for $N = 38$, $75-77$ should also be included.

The most successful methods applied to the global optimization of LJ clusters are the Basin-Hopping algorithm⁷ (based on Monte Carlo with minimization^{18,19}) and genetic algorithms.⁶ The Basin-Hopping algorithm has been able to locate all known lowest-energy minima of LJ clusters up to 110 atoms (including those for $N = 38$ and $75-77$), some of which were never found previously. The continuous genetic algorithm⁶ found new lowest-energy minima for LJ₈₈ and LJ₉₈; however, it was not able to find global minima for $N = 75-78$ but found previously known second-to-the-lowest energy structures instead.

A promising approach to surmount the multiple-minima problem involves methods based on the deformation of the original rugged energy surface, thereby reducing the number of minima by orders of magnitude, at best even to a single minimum, and simplifying the conformational search greatly.^{10,21-29} The simplest approach to deformation-based global optimization is to track the deformed structure back from the lowest-energy

minimum on the highly deformed PES to the undeformed PES (reversing procedure); however, it is successful only for relatively simple systems. Usually, the lowest-energy minimum on the highly deformed PES does not correspond directly to the global minimum of the original PES (even when there is only one minimum for the highest deformation), and a single trajectory connecting the highly deformed and undeformed minima often branches during the reversal of the deformation. A possible solution to this problem is to track back more than one minimum and try to detect branching of a trajectory by using a local search in the vicinity of a trajectory. This approach was found to be more successful than the single or multiple-trajectory approach and was applied in the theoretical prediction of crystal structures.^{30,31} However, it does not work for highly demanding applications, such as large Lennard-Jones clusters or polypeptide chains.

In the present paper, a recently proposed self-consistent basin-to-deformed-basin mapping (SCBDBM)³² method is used. Its underlying principle is to locate a group of large basins containing low-energy minima (superbasins) in the original energy surface. This is achieved by coupling the superbasins in the original PES to basins in a highly deformed energy surface by iterative cycles, each consisting of reversing the deformation and then deforming the newly found low-energy structures. This method has already been applied successfully to predict the lowest-energy structures of polyalanine chains of length up to 100 amino acid residues.³²

2. Method

2.1. Global Optimization Algorithm. Details of the method are presented elsewhere;³² here, we present only a brief description. We consider a function $f(\mathbf{x})$ of several variables. Consider a mapping $F(\mathbf{x}, a)$, where a defines the extent of deformation, such that $F(\mathbf{x}, a)$ becomes smoother with a gradually smaller number of energy minima with increasing a ; assume that $F(\mathbf{x}, 0) = f(\mathbf{x})$. In order to utilize the deformation F to locate the global minimum of f , we first note that, with increasing a , the number of minima gradually decreases because some of the minima merge into one. As deformation proceeds, groups of individual minima are first merged into single minima, defining the *superbasins* of these groups of minima for certain values of the deformation parameter. As the deformation parameter a increases, the minima continue to merge, causing the superbasins from smaller deformations (*lower-order superbasins*) to merge, constituting higher-order superbasins (for larger deformation). Finally, for a very high deformation, only a few minima remain. A logical procedure to find the global minimum of $f(\mathbf{x})$ would, therefore, be first to locate the highest-order (the most deformed) superbasin in $F(\mathbf{x}, a)$ related to this minimum and then to locate within it the superbasins of gradually lower order (less deformed) that still contain this minimum, until the deformation is fully reversed, i.e., the global minimum in the original energy surface is located. However, the major difficulty in proceeding in this manner is that there is no straightforward relation between the values of F at its minima and the corresponding minimum values of f . Therefore, one can never tell which superbasin corresponds to the global minimum of the original energy function, based only on the "energy" relations between superbasins. As a consequence, it is clearly insufficient to reverse the deformation only once, in order to find the global minimum of f , even if a multitrajectory search is carried out during the reversing procedure. To surmount this problem, we have proposed³² a self-consistent procedure that finds the coupling between the superbasins of different order, which is achieved

TABLE 1: Results of Global Optimization with the SCBDBM Method for Lennard-Jones Clusters

no. of atoms	energy, reduced units			
	global minimum	random search ^a	SCBDBM	SCBDBM (no MCM) ^b
5	-9.103 852	-9.103 852 (62.3)	-9.103 852	-9.103 852
6	-12.712 062	-12.712 062 (66.0)	-12.712 062	-12.712 062
7	-16.505 384	-16.505 384 (13.5)	-16.505 384	-16.505 384
8	-19.821 489	-19.821 489 (34.8)	-19.821 489	-19.821 489
9	-24.113 360	-24.113 360 (14.0)	-24.113 360	-24.113 360
10	-28.422 532	-28.422 532 (4.1)	-28.422 532	-28.422 532
11	-32.765 970	-32.765 970 (3.1)	-32.765 970	-32.765 970
12	-37.967 600	-37.967 600 (2.5)	-37.967 600	-37.967 600
13	-44.326 801	-44.326 801 (1.3)	-44.326 801	-44.326 801
14	-47.845 157	-47.845 157 (3.8)	-47.845 157	-47.845 157
15	-52.322 627	-52.322 627 (3.0)	-52.322 627	-52.322 627
16	-56.815 742	-56.815 742 (1.8)	-56.815 742	-56.815 742
17	-61.317 995	-61.317 995 (0.7)	-61.317 995	-61.317 995
18	-66.530 949	-66.530 949 (0.1)	-66.530 949	-66.530 949
19	-72.659 782	-72.659 782 (0.2)	-72.659 782	-72.659 782
20	-77.177 043	-77.177 043 (0.8)	-77.177 043	-77.177 043
21	-81.684 571	-81.684 571 (0.1)	-81.684 571	-81.684 571
22	-86.809 782	-86.809 782 (0.2)	-86.809 782	-86.809 782
23	-92.844 472	-92.844 472 (0.1)	-92.844 472	-92.844 472
24	-97.348 815	-97.348 815 (0.1)	-97.348 815	-97.348 815
25	-102.372 663	-102.372 663 (0.1/0.12)	-102.372 663	-102.372 663
26	-108.315 616	-106.998 484 (0.0/0.01)	-108.315 616	-108.315 616
27	-112.873 584	-112.873 584 (0.3/0.21)	-112.873 584	-112.873 584
28	-117.822 402	-117.822 402 (0.1/0.07)	-117.822 402	-117.822 402
29	-123.587 371	-123.587 371 (0.1/0.02)	-123.587 371	-123.587 371
30	-128.286 571	-127.761 894 ^c (0.0/0.00)	-128.286 571	-128.286 571
31	-133.586 422	-132.247 370 (0.0/0.01)	-133.586 422	-133.586 422
32	-139.635 524	-137.869 586 (0.0/0.01)	-139.635 524	-139.635 524
33	-144.842 719	-143.380 151 (0.0/0.01)	-144.842 719	-144.842 719
34	-150.044 528	-148.547 367 ^d (0.0/0.00)	-150.044 528	-150.044 528
35	-155.756 643	-153.392 605 ^e (0.0/0.00)	-155.756 643	-155.756 643
36	-161.825 363	-158.676 188 (0.0)	-161.825 363	-161.825 363
37	-167.033 672	-167.017 283 (0.0)	-167.033 672	-167.033 672
38	-173.928 427	-171.147 501 (0.0)	-173.928 427	-173.928 427
39	-180.033 185	-175.983 755 (0.0)	-180.033 185	-180.033 185
40	-185.249 839	-181.603 835 (0.0)	-185.249 839	-185.249 839
41	-190.536 277	-187.664 301 (0.0)	-190.536 277	-190.536 277
42	-196.277 534	-191.703 703 (0.0)	-196.277 534	-196.277 534
43	-202.364 664	-198.218 052 (0.0)	-202.364 664	-202.364 664
44	-207.688 728	-206.289 110 (0.0)	-207.688 728	-207.688 728
45	-213.784 862	-209.609 291 (0.0)	-213.784 862	-213.784 862
46	-220.680 330	-214.417 990 (0.0)	-220.680 330	-220.680 330
47	-226.012 256	-224.433 093 (0.0)	-226.012 256	-226.012 256
48	-232.199 529	-227.501 771 (0.0)	-232.199 529	-232.199 529
49	-239.091 864	-232.721 425 (0.0)	-239.091 864	-239.091 864
50	-244.549 926	-239.468 963 (0.0)	-244.549 926	-244.549 926
51	-251.253 964	-245.204 671 (0.0)	-251.253 964	-251.253 964
52	-258.229 991	-250.545 977 (0.0)	-258.229 991	-258.229 991
53	-265.203 016	-262.141 953 (0.0)	-265.203 016	-265.203 016
54	-272.208 631	-264.672 834 (0.0)	-272.208 631	-272.208 631
55	-279.248 470	-268.622 886 (0.0)	-279.248 470	-279.248 470
56	-283.643 105		-283.643 105	
57	-288.342 625		-288.342 625	
58	-294.378 148		-294.378 148	
59	-185.249 839		-299.738 070	
60	-305.875 476	-300.131 608 (0.0)	-305.875 476	-305.875 476
61	-312.008 896		-312.008 896	
62	-317.353 901		-317.353 901	
63	-323.489 734		-323.489 734	
64	-329.620 147		-329.620 147	
65	-334.971 532	-328.870 300 (0.0)	-334.971 532	-334.971 532
66	-341.110 599		-341.110 599	
67	-347.252 007		-347.252 007	
68	-353.394 542		-353.394 542	
69	-359.882 566		-359.882 566	
70	-366.892 251	-356.097 931 (0.0)	-366.892 251	-366.313 126
71	-373.349 661		-373.349 661	
72	-378.637 253		-378.637 253	
73	-384.789 377		-384.789 377	
74	-390.908 500		-390.908 500	
75	-397.492 331	-386.545 271 (0.0)	-396.282 249	-396.238 512

TABLE 1: Continued

no. of atoms	energy, reduced units			
	global minimum	random search ^a	SCBDBM	SCBDBM (no MCM) ^b
76	-402.894 866		-402.384 580	
77	-409.083 517		-408.518 265	
78	-414.794 401		-414.794 401	
79	-421.810 897		-421.810 897	
80	-428.083 564	-419.885 749 (0.0)	-428.083 564	-427.080 532
81	-434.343 643		-434.343 643	
82	-440.550 425		-440.550 425	
83	-446.924 094		-446.924 094	
84	-452.657 214		-452.657 214	
85	-459.055 799	-449.427 020 (0.0)	-459.055 799	-458.283 317
86	-465.384 493		-465.384 493	
87	-472.098 165		-472.098 165	
88	-479.032 630		-479.032 630	
89	-486.053 911		-486.053 911	
90	-492.433 908	-477.912 316 (0.0)	-492.433 908	-489.091 398
91	-498.811 060		-498.811 060	
92	-505.185 309		-505.185 309	
93	-510.877 688		-510.877 688	
94	-517.264 131		-517.264 131	
95	-523.640 211	-514.356 127 (0.0)	-523.640 211	-520.731 120
96	-529.879 146		-529.879 146	
97	-536.681 383		-536.681 383	
98	-543.642 957		-543.642 957	
99	-550.666 526		-550.666 526	
100	-557.039 820	-542.012 683 (0.0)	-557.039 820	-554.634 820

^a The lowest-energy minimum found in 1000 local minimizations starting from randomly generated cluster configurations. The probability (in percent) of finding the global minimum in 1000 random searches is reported in parentheses; the probability of finding the global minimum in 10 000 random searches is reported as the second number in parentheses when appropriate. Even probabilities as small as 0.1 or 0.01, for 1000 and 10 000 random searches, respectively, are large enough for successful identification of global minima. ^b No MCM is carried out on the undeformed surface. ^c The lowest-energy minimum found in 10 000 random searches was -128.181 578. ^d The lowest-energy minimum found in 10 000 random searches was -149.996 978. ^e The lowest-energy minimum found in 10 000 random searches was -154.804 752.

by iterating the steps consisting of reversing the deformation and, subsequently, reintroducing the deformation.

The procedure is outlined in Figure 1. It consists of a series of *macroiterations*. Each macroiteration establishes the coupling between superbins of consecutive order (and contains a self-consistent procedure within it). In macroiteration i , the parameter $a^{(i)}$, which controls the deformation, changes between two extreme values $a_{\max}^{(i)}$ and $a_{\min}^{(i)}$. For macroiteration $i + 1$, $a_{\max}^{(i+1)} = a_{\min}^{(i)}$ and $a_{\min}^{(i+1)} = a_{\min}^{(i)}/\Delta$ (or 0 in the last macroiteration), where Δ is a logarithmic step length. The first macroiteration is initialized with randomly generated and minimized conformations in the highest deformed space, while each subsequent macroiteration is fed with the results of the previous one.

Within each macroiteration, minima found on the highest deformed surface ($a_{\max}^{(i)}$) are tracked back to the least deformed surface ($a_{\min}^{(i)}$) by decreasing the deformation parameter a and searching locally for new minima at each reversal step. This is called the reversing procedure. Subsequently, new minima found on the least deformed surface are tracked to the highest deformed surface by gradually increasing the parameter a and performing local minimization (without any local search) at each step. This is the reversed-reversing procedure. The maximum number of trajectories to be followed is fixed at p . The iteration described above (one reversing procedure and one reversed-reversing procedure) is iterated until no new minima are found in the least deformed surface or the maximum allowed number of iterations is exceeded.

The local search plays a very important role in the algorithm by detecting branching of minima during the reversing procedure. However, this search should be carried out in the vicinity of a starting minimum; otherwise the relationship between

minima may be lost (i.e., the newly found minimum may not be related to the previous one, but belong to a completely different "tree" of trajectories). The simplest local search is a random perturbation of a structure followed by a local energy minimization, as implemented in the multiple-trajectory perturbation approach.^{30,31} During a perturbation, the positions of all atoms in a cluster are perturbed randomly (the amplitude being lower than a predefined number δ); if the perturbation is very small the search is local, but quite often the subsequent energy minimization simply restores the original structure. On the other hand, if the perturbation is larger, the repulsion between perturbed atoms in the core of a cluster usually forces the local minimization procedure to jump quite far from the starting point (and the cluster temporarily explodes during local minimization); even relatively small perturbations of the core atoms frequently resulted in a nonlocal search. The solution is simply to make the size of a perturbation for a given atom depend linearly on the distance from the center of a cluster to that atom, with the atom at the center of a cluster not being perturbed at all, and the surface atoms being perturbed to the maximum extent δ . In this approach, all atoms in a cluster are still perturbed (and, therefore, changes in the core of the cluster are possible), but the search remains local.

Another version of local search used in our algorithm is a linear search along randomly generated directions in the multidimensional space; the search stops when a new basin is found or a predefined maximum number of steps is exhausted. This kind of search nearly always finds neighboring basins and rarely stays in the starting basin.

We have achieved the best results (i.e., the fastest convergence to the global minimum) when both kinds of local searches described above were used in our algorithm. The local search

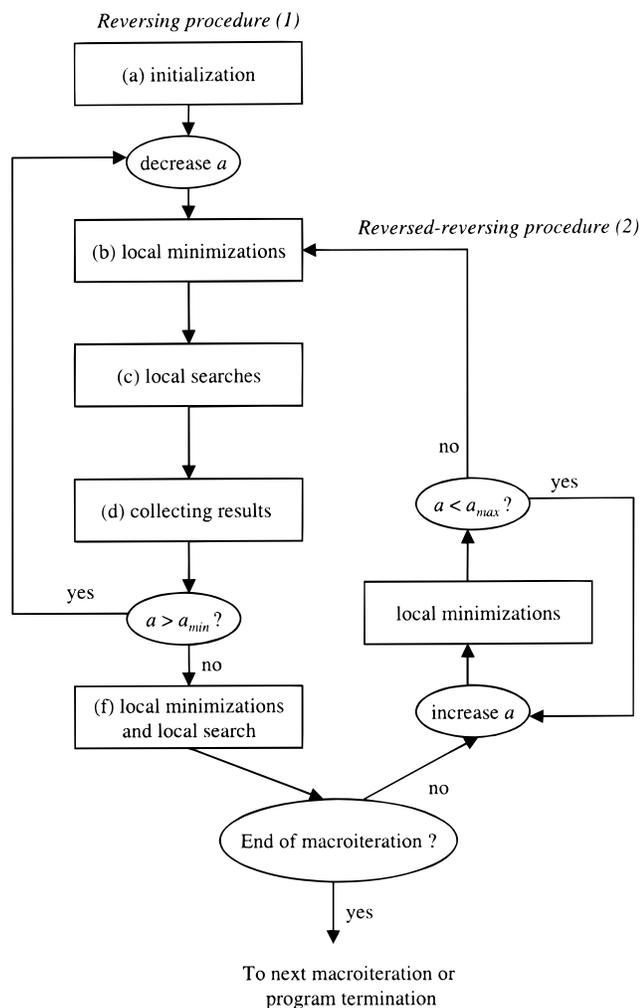


Figure 1. Block diagram of the reversing procedure coupled with the reversed-reversing procedure within a single macroiteration.

on the undeformed energy surface is treated differently if $a_{\min} = 0$. In this case, after every predefined number of iterations within a macroiteration (usually 4), a short Monte Carlo with minimization (MCM)^{18,19} search is carried out (until 50 structures are accepted by the Metropolis criterion), instead of using the local search procedures described above.

2.2. Deformation. A minimal requirement for deformation to be useful in the SCBDBM method is that the number of minima of a potential function should decrease when the deformation increases, causing local minima to merge. Merging of two local minima may occur in two different ways: first, two potential wells overlap continuously as the deformation increases, so that there is no point at which one of them vanishes—all three critical points (two minima and a saddle point) simultaneously turn into a new minimum; second, one of the minima vanishes (the minimum and a saddle point turn into an inflection point), and the local minimization procedure starting from the present point finds the second minimum. In the second case, one minimum may vanish geometrically close to the other minimum. In such a situation, a local search may be able to find this vanishing minimum during the reversing procedure; otherwise, when the first minimum vanishes geometrically far from the first one, there is only a remote chance of finding the former minimum. Therefore, the deformation should promote a continuous merging of minima, or at least this kind of merging, in which the minima are geometrically close.

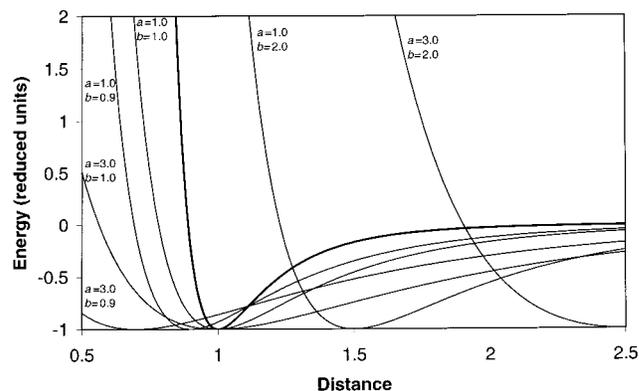


Figure 2. Application of the DSM to the Lennard-Jones potential. Thick solid line shows the original LJ potential. Other lines show progressing deformation for the values of a and b indicated.

The distance scaling method (DSM) has been chosen as the deformation procedure for the present work. It is very simple to implement, is designed to work with pairwise interactions, and was shown to perform well in finding the global minima of Lennard-Jones and water clusters,^{10,24,25} and in predicting the crystal structures of small molecules.³¹ In the DSM,¹⁰ the site–site distance r_{ij} is transformed to \tilde{r}_{ij} as follows:

$$\tilde{r}_{ij} = \frac{r_{ij} + ar_{o,ij}}{1 + ba} \quad (2)$$

The parameter $r_{o,ij}$ in eq 2 is the position of the minimum in the undeformed pairwise-interaction function under consideration. On increasing the deformation parameter a , the original function of the site–site distance (e.g., the Lennard-Jones potential) is flattened, but the position of its minimum and the function value at the minimum remain the same, if the value of the parameter b is taken as 1 (as in the original formulation¹⁰ of the DSM). The parameter b controls the position of the minimum and remains constant during the calculations. A value of $b > 1$ means that the position of the minimum of the deformed site–site function will shift to larger values, while for $b < 1$ it will shift toward zero, and a two-body potential will become totally attractive for $a = 1/(1 - b)$. For $b = 0$, the deformation works by shifting the potential toward zero distance; therefore, the repulsion part of the potential function is cut, and the potential becomes purely attractive for $a = 1$.

Application of the deformation to the Lennard-Jones pairwise potential makes it relatively long-ranged while diminishing energy barriers between minima due to lowering repulsion (for all values of b) and attraction (for $b > 0$). The reason for this behavior is that the flattened potential converges to zero more slowly with increasing distance. Flattenning and shifting of the deformed LJ potential for three different values of b (0, 1, and 2) is illustrated in Figure 2.

In order to test the influence of different values of the parameter b on the merging of local minima during the deformation, the reversed-reversing procedure was carried out for the 100 lowest-energy minima of the LJ₃₈ cluster for different values of b . The resulting trajectories of all 100 minima for $b = 2, 1$, and 0 are shown in Figures 3, 4, and 5, respectively. The plots show the relative deformed energies of the structures (relative to the lowest deformed energy found for a given value of the deformation parameter a). Rather than plotting the deformed energies themselves, the relative values better illustrate the influence of the deformation on energy ordering and merging of minima because the energy of a cluster changes while the deformation increases; the pairwise potential energy becomes

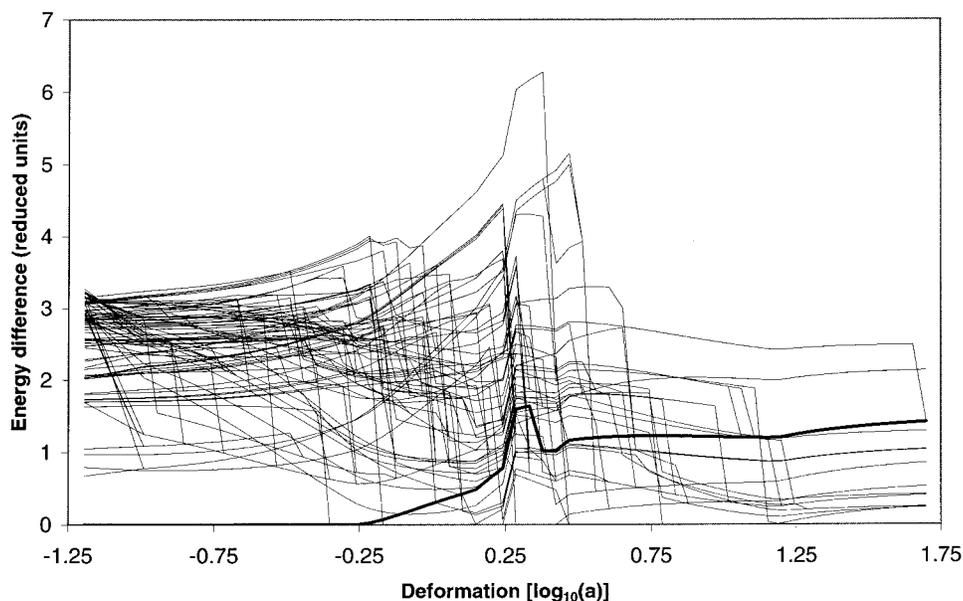


Figure 3. Changes in the order of the energies during the reversed-reversing procedure for the 100 low-energy minima in the undeformed energy surface of LJ_{38} for the parameter $b = 2.0$. The vertical axis describes the energy difference between the current minimum and the lowest-energy minimum for a given value of the deformation parameter. The global minimum trajectory is shown with a thick solid line. The procedure starts from the 100 minima in the undeformed energy surface ($a \sim 0$). With increasing deformation, the trajectories ultimately merge to around 10 minima near $a = 50$ ($\log a = 1.7$). It can be seen that the global minimum trajectory appears at a higher energy at $a = 50$, even though it was the lowest energy in the undeformed surface. Further, despite the merging, there are still many left at $a = 50$.

flatter, and the total energy of a cluster (being a sum of all pairwise interactions) becomes significantly lower for higher deformation. Therefore, the differences between the deformed energies can hardly be noticed on the graph because of the much more significant decrease of the deformed energy with increasing deformation. Abrupt vertical drops in the energy difference leading to connecting of two different lines, present in all three plots, indicate merging of minima.

For $b = 2.0$ (Figure 3), the merging of minima is distributed uniformly throughout the deformation range when using a logarithmic scale for the deformation. The order of the minima changes as the deformation increases, and the global minimum (i.e., for $a = 0$) becomes a local minimum for deformations larger than $a = 0.5$. During the reversed-reversing procedure, 90% of the minima merged, leaving only 10 different basins for the maximum deformation, and the remaining minima did not merge for reasonably large deformations (up to $a = 100$). It should be noted that, if the deformation a increases too much, then the resulting surfaces are extremely flat, and the local minimization procedure often fails. The value 2.0 of the parameter b shifts the local minimum of the LJ potential for each pair of interacting atoms to the right (larger distances), causing the cluster to grow in size. As a result of such growth, only minima that are geometrically close merge; the other minima move apart from each other faster than deformation is able to lower the barriers between them. This behavior of the DSM is qualitatively the same as the behavior of the diffusion equation method (DEM)²² applied directly to a Gaussian approximation of the Lennard-Jones potential. This value of b is, therefore, not an appropriate choice for deformation of LJ clusters. For $b = 1.0$ (Figure 4) 98% of all minima merged for high deformation, leaving only two local minima for $a = 30$ ($\log a = 1.48$). Unfortunately, most merging occurs in a small region of the deformation (around $a = 1$), even when using a logarithmic scale for the deformation. In this case, detection of new minima during the reversing procedure may be difficult, because many new minima appear at the same stage of the

deformation; i.e., it is difficult to pick starting points at $a = 1$ with which to carry out the reversing procedure.

The third, extremal, value of $b = 0$ (Figure 5) seems to work best. For this value, the most appropriate scale of deformation is linear rather than logarithmic, and the minima merge uniformly throughout the whole deformation range, allowing much better search for higher deformation. A minimum at high deformation usually corresponds to many minima in the undeformed surface. Therefore, if a minimum is not detected in deformed surface, then a whole family of corresponding minima in the undeformed surface is lost. There is only one minimum left for the deformation parameter $a = 1$, because of the purely attractive behavior of the potential at this stage of deformation. The order of minima changes during the deformation, indicating the necessity of using the energies of only undeformed structures when deciding which minima to store during the reversing procedure, as described in section 2.1 of ref 30. The differences between the deformed energies of nonmerging minima increase for the deformation parameter $a < 0.5$, and then decrease for $a > 0.5$, which divides the range of the deformation parameter into two parts; therefore, we have chosen to use two macroiterations with $a_{\max}^1 = 1$, $a_{\min}^1 = 0.5$, and with $a_{\max}^2 = 0.5$ with $a_{\min}^2 = 0$.

3. Results and Discussion

All calculations were carried out using the numerical parameter controlling the deformation described in the previous section, viz., $b = 0$. The number of trajectories was chosen as 5, and the number of macroiterations as 2; for four different values of the deformation a , a local search was carried out during the reversing procedure. The maximum deformation parameter a_{\max} was 1.0 (with b held constant at the value 0), and the deformation parameter a was changed linearly during the reversing procedure. The maximum number of iterations within the first macroiteration (one iteration being a pair of reversed and reversed-reversing procedures) was set as 4, the maximum number of iterations within the second macroiteration was set

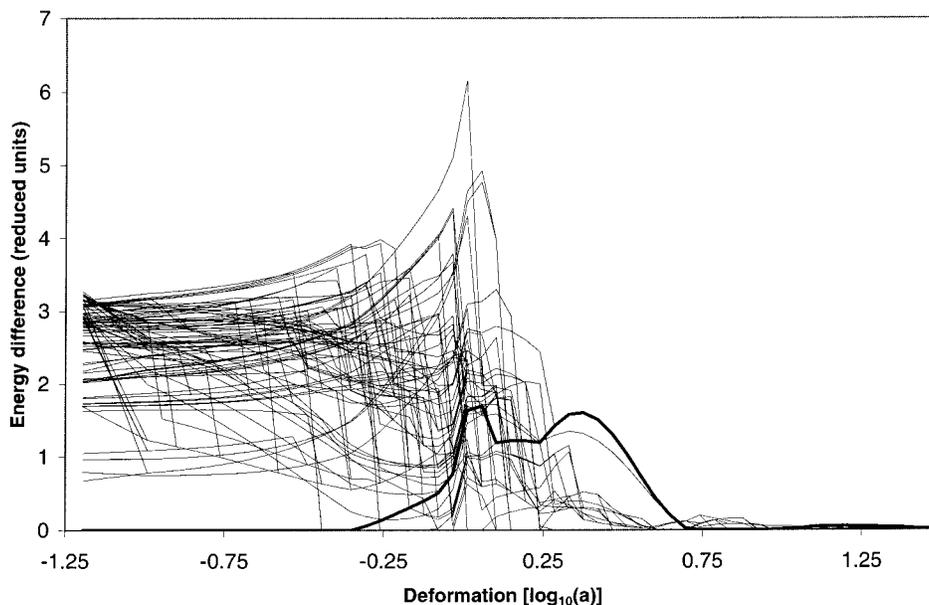


Figure 4. Same as Figure 3, but for $b = 1.0$. In contrast to Figure 3, all trajectories have merged to only two minima at $a = 30$ in the deformed surface. All the merging of the trajectories has taken place in the vicinity of $a = 1$.

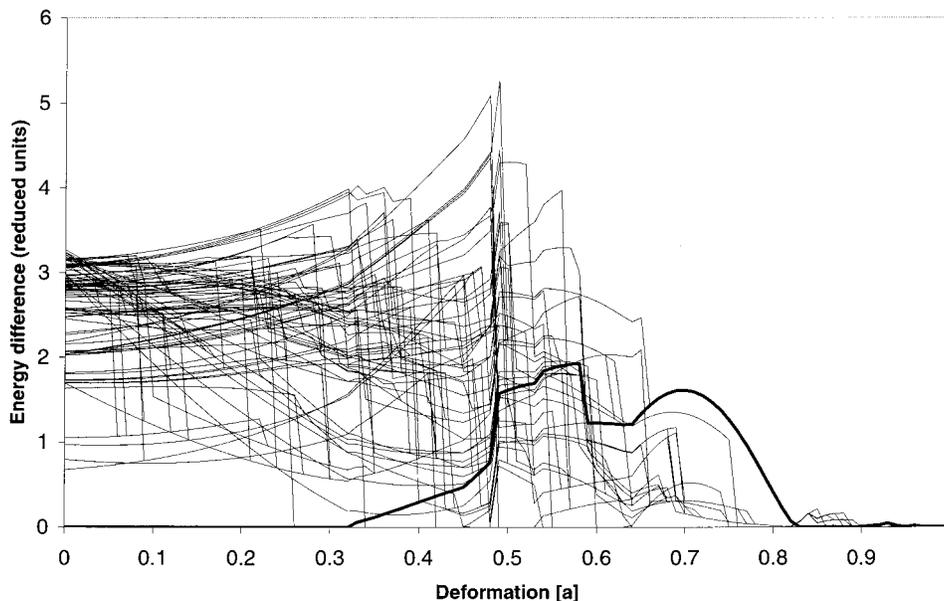


Figure 5. Same as Figure 4, but for $b = 0.0$. In contrast to Figure 4, there is only one minimum at $a = 1$, and the trajectories merge more uniformly in the deformation dimension, i.e., there are many trajectories in the region of $a > 0.5$ whereas, in Figure 4, the right-half of the plot is devoid of trajectories.

at 11. The program has been parallelized on a coarse grain level, i.e., local searches and trajectory tracking were carried out in parallel. All calculations were carried out on an IBM SP2 supercomputer at the Cornell Theory Center, and the resources consumed varied greatly with the cluster size. For LJ₇₀, full global optimization required 3.5 h, using 10 processors of the SP2 supercomputer. The typical number of local minimizations carried out for larger clusters was about 40 000. It must be noted, however, that most of these minimizations were used to determine trajectories in the reversing and reversed-reversing procedures. In this case, a minimization was started from the previously minimized structure at a slightly different deformation parameter (slightly larger a for the reversing procedure, slightly smaller a for the reversed-reversing procedure), and, therefore is extremely fast. Most of the numerical expense is involved in carrying out local minimizations in the local search;

the typical number of local searches carried out for larger clusters was 8000.

In order to validate the parameters used, full global-optimization runs with the numerical effort doubled (by doubling the number of iterations within the macroiteration) were carried out for 10 clusters, whose sizes were chosen randomly between 50 and 100. The SCBDBM results were completely independent of the starting geometries; different global optimization runs starting from different randomly generated initialization parameters (seeds) resulted in exactly the same lowest-energy minima.

The detailed results are shown in Table 1. For comparison, the results of random searches (column 3) and the SCBDBM method without the MCM search on the undeformed surface (last column) are provided. The random search consisted of 1000 local minimizations, starting from randomly generated structures; in order to obtain random starting points, the Cartesian

coordinates of all atoms in the cluster were selected randomly as a number from the [0,8] range (in reduced units). The lowest-energy minimum found in a random search is reported for each cluster size along with the probability of finding the global minimum in such a search (the same column, with the percent probability in parentheses). For clusters LJ_N with $N = 25-35$, an additional 9000 local minimizations starting from randomly selected structures were carried out, and the overall probability of finding the global minimum for these clusters in 10 000 random local minimizations is reported as the second number in parentheses.

No previously unknown lowest-energy cluster geometries were found by the SCBDBM method, and the method correctly located the currently known global minima for all clusters, except for LJ_{75} , LJ_{76} , and LJ_{77} . In these cases, the previously known lowest-energy icosahedral structures (being the second to the lowest in energy) were located. It has been shown that long-range pairwise potentials promote icosahedral structures over fcc structures in atomic clusters.³³ When the deformation is increased, the potential becomes long range, which probably makes most nonicosahedral (fcc) structures higher in energy, and, finally eliminates them. However, in the case of LJ_{38} , despite the fact that the fcc structure (the global minimum for $a = 0$) becomes higher in the deformed energy for the intermediate values of the deformation parameter a , it remains separated, and merges geometrically close to the icosahedral structure for a very high value of the deformation parameter (see Figure 5). Unfortunately, the basins corresponding to nonicosahedral structures for LJ_{75-77} disappear geometrically very far from the main group of trajectories (because of the long-range deformed potential), making them virtually impossible to find. In order to improve the results, a better deformation procedure should be designed, where the nonicosahedral minima would merge geometrically close to the icosahedral ones.

The SCBDBM method carried out without the MCM search on the undeformed surface (last column of Table 1) is successful up to LJ_{65} , as indicated by the deviation in the energy beyond 65 atoms. This shows that the DSM coupled with a local search and self-consistent mapping is a very powerful global optimization method. Its effectiveness could be increased further by increasing the number of local searches and the limit of the number of iterations within a macroiteration; however, the numerical cost of such change is high. A much more economical way to improve the results is to add a short MCM search in the vicinity of the minima in the undeformed surface, as we have done here.

The results of the random search (third column in Table 1) show that the task of finding the global minima of LJ clusters of sizes smaller than 30 atoms is relatively easy; the probability of randomly finding the global minimum is about 0.1% or larger for these clusters (see footnote a in Table 1). Finding the global minima for clusters LJ_{30-35} is, in turn, of moderate difficulty.

The interesting question about the SCBDBM method is how important is the deformation in the performance of the method. To answer this question, full global optimization runs with the deformation parameter a set permanently to zero were carried out for different clusters. The results (not presented here) showed that SCBDBM without the deformation works successfully (with some exceptions) up to about 45 atoms in a cluster, which means

that it performs better than the corresponding random search for clusters containing about 10 atoms more than the number used in successful random searches.

Based on the present work, the SCBDBM method has been shown to be a very efficient method of global optimization, comparable with continuous genetic algorithms or the Basin-Hopping algorithm, at least in applications to Lennard-Jones clusters.

Acknowledgment. This research was supported by grants from the National Science Foundation (MCB95-13167), and from the National Institutes of Health (GM-14312). The computations were carried out at the Cornell Theory Center which is funded in part by NSF, New York State, the NIH National Center for Research Resources (P41RR-04293), the IBM Corp., and the CTC Corporate Partnership Program.

References and Notes

- (1) Anfinsen, C. B. *Science* **1973**, *181*, 223-230.
- (2) Wales, D. J.; Scheraga, H. A. *Science* **1999**, *285*, 1368-1372.
- (3) Scheraga, H. A.; Lee, J.; Pillardy, J.; Ye, Y.-J.; Liwo, A.; Ripoll, D. R. *J. Global Optimization*, in press.
- (4) Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*, 11227-11237.
- (5) Northby, J. A. *J. Chem. Phys.* **1987**, *87*, 6166.
- (6) Deaven, D. M.; Tit, N.; Morris, J. R.; Ho, K. M. *Chem. Phys. Lett.* **1996**, *256*, 195.
- (7) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111-5116.
- (8) Wales, D. J. The Cambridge Cluster Database; URL <http://brian.ch.cam.ac.uk>.
- (9) Mackay, A. L. *Acta Crystallogr.* **1962**, *15*, 916.
- (10) Pillardy, J.; Piela, L. *J. Phys. Chem.* **1995**, *99*, 11805-11812.
- (11) Doye, J. P. K.; Wales, D. J.; Berry, R. S. *J. Chem. Phys.* **1995**, *103*, 4234.
- (12) Doye, J. P. K.; Wales, D. J. *Chem. Phys. Lett.* **1995**, *247*, 339.
- (13) Marks, L. D. *Philos. Mag. A* **1981**, *49*, 81.
- (14) van de Waal, B. W. *J. Cryst. Growth* **1995**, 153-155.
- (15) van de Waal, B. W. *Phys. Rev. Lett.* **1996**, 1083-1086.
- (16) Doye, J. P. K.; Wales, D. J.; Miller, M. A. *J. Chem. Phys.* **1998**, *109*, 8143-8153.
- (17) Wales, D. J.; Miller, M. A.; Walsh, T. R. *Nature* **1998**, *394*, 758.
- (18) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci., U.S.A.* **1987**, *84*, 6611-6615.
- (19) Li, Z.; Scheraga, H. A. *J. Mol. Struct. (THEOCHEM)* **1988**, *179*, 333-352.
- (20) Goldberg, D. E. *Genetic Algorithms in Search, Optimization & Machine Learning*; Addison-Wesley: Reading, MA, 1989.
- (21) Piela, L.; Kostrowicki, J.; Scheraga, H. A. *J. Phys. Chem.* **1989**, *93*, 3339-3346.
- (22) Kostrowicki, J.; Piela, L.; Cherayil, B. J.; Scheraga, H. A. *J. Phys. Chem.* **1991**, *95*, 4113-4119.
- (23) Kostrowicki, J.; Piela, L. *J. Optimiz. Theory App.* **1991**, *69*, 269-284.
- (24) Pillardy, J.; Olszewski, K. A.; Piela, L. *J. Phys. Chem.* **1992**, *96*, 4337-4341.
- (25) Pillardy, J.; Olszewski, K. A.; Piela, L. *J. Mol. Struct. (THEOCHEM)* **1992**, *270*, 277-285.
- (26) Amara, P.; Hsu, D.; Straub, J. E. *J. Phys. Chem.* **1993**, *97*, 6715-6721.
- (27) Oresić, M.; Shalloway, D. *J. Chem. Phys.* **1994**, *101*, 9844-9857.
- (28) Ma, J. P.; Straub, J. E. *J. Chem. Phys.* **1994**, *101*, 533.
- (29) Pappu, R. V.; Hart, R. K.; Ponder, J. W. *J. Phys. Chem. B* **1998**, *102*, 9725-9742.
- (30) Wawak, R. J.; Gibson, K. D.; Liwo, A.; Scheraga, H. A. *Proc. Natl. Acad. Sci., U.S.A.* **1996**, *93*, 1743-1746.
- (31) Wawak, R. J.; Pillardy, J.; Liwo, A.; Gibson, K. D.; Scheraga, H. A. *J. Phys. Chem. A* **1998**, *102*, 2904-2918.
- (32) Pillardy, J.; Liwo, A.; Groth, M.; Scheraga, H. A. *J. Phys. Chem. B* **1999**, *103*, 7353-7366.
- (33) Doye, J. P. K.; Wales, D. J. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 4233-4243.