INFORMS Journal on Computing
Vol. 16, No. 4, Fall 2004, pp. 371–379
ISSN 0899-1499 | EISSN 1526-5528 | 04 | 1604 | 0371

DOI 10.1287/ijoc.1040.0084

Global Optimization of Morse Clusters by Potential Energy Transformations

Jonathan P. K. Doye
University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom, jpkd1@cam.ac.uk

Robert H. Leary
San Diego Supercomputer Center, University of California, San Diego, San Diego, California 92186-9784, USA, leary@sdsc.edu

Marco Locatelli
Dipartimento di Informatica, Università di Torino, Corso Svizzera 185, I-10149 Torino, Italy, locatell@di.unito.it

Fabio Schoen
Dipartimento di Sistemi e Informatica, Università di Firenze, Via di S.Marta, 3, I-50139 Firenze 185, Italy, schoen@ing.unifi.it

The Morse potential is a simple model for the potential energy of atoms with a single parameter ρ that determines the width of the potential well and allows a wide variety of materials to be modeled. Morse clusters are particularly important for applications, but their global optimization is also an extremely hard problem, highly relevant to methods that are to be applied to find the optimal configuration of a biomolecule. In particular, large ρ values are very challenging and, until now, no unbiased global-optimization method has been able to detect all the (putative) global minima at $\rho = 14$ for clusters with up to $N = 80$ atoms. In this paper we introduce some techniques for transforming the original Morse potential that allow us to increase considerably the efficiency in locating the known global minima and also to discover some new optimal clusters. These methods are promising candidates for application to the optimization of biomolecules.

Key words: Morse potential; global optimization; basin-hopping; two-phase local search
History: Accepted by Harvey J. Greenberg, Guest Editor; received July 2003; revised February 2004; accepted March 2004.

1. Introduction

Lennard-Jones clusters have become an archetypal test system for global optimization algorithms designed to find the lowest energy configuration of a model "molecular" system, be it a cluster, a biomolecule, or a crystal (Wales and Scheraga 1999). Hundreds of published papers have been dedicated to various aspects of the Lennard-Jones cluster problem. There probably is no equivalent test system for the protein-folding problem, at least when modeled in a continuous configuration space, and many of the most successful algorithms for Lennard-Jones clusters have been subsequently applied to optimize polypeptides and proteins. For proteins modeled on a lattice the archetypal test system is probably the HP model (Lau and Dill 1989). However, global-optimization techniques for systems in discrete and continuous spaces have substantial differences.

The choice of Lennard-Jones clusters is due to a number of reasons. The Lennard-Jones potential is extremely simple. Excellent putative global minima have now been tabulated for clusters with up to 309 atoms (Romero et al. 1999, Wales and Doye 1997). Lennard-Jones clusters have also become a model system for understanding a wide range of cluster properties, such as their structure, thermodynamics, dynamics, and energy landscapes, allowing the ease or difficulty of global optimization to be correlated with these cluster properties (Doye and Wales 1998).

However, Lennard-Jones clusters have a number of deficiencies as a test system. First, they are structurally rather uniform. Virtually all of the global minima in the relevant size range are based on Mackay icosahedra (Northby 1987). Second, for systems of their dimensionality, Lennard-Jones clusters are relatively easy to optimize. This is because the potential energy landscapes for these systems give rise to strong funneling properties, directing the system down towards the global minimum (Doye et al. 1999). For example, even though there are an estimated 10^{21} geometrically distinct minima for a 55-atom Lennard-Jones cluster, the basin-hopping algorithm is able to find the global minimum from a random starting point after only sampling on average 150 of those minima (Wales and Scheraga 1999).

The most difficult Lennard-Jones clusters to optimize occur at the few sizes where the global minimum has a structure that is not in fact based on the Mackay icosahedron. In these cases, the energy landscape has a double-funnel topography, where the dominant funnel leads down to icosahedral structures and there is a second narrower, and less accessible, funnel leading down to the global minimum (Doye et al. 1999). Typically, these examples are one or two orders of magnitude more difficult (Leary 2000). The fact that a few examples are the most meaningful tests of the performance of an algorithm is probably not the behavior one wants from a test system. Rather, it is much better to have more uniform difficulty.

Because of these deficiencies, Morse clusters have been suggested as a much more rigorous alternative both as a more challenging test system, and as a more accurate model for understanding the structure of real-life clusters (Doye and Wales 1997). The Morse (1929) pair potential as a function of the interatomic distance r between two atoms is defined as

$$v_\rho(r) = [\exp\{\rho(1-r)\} - 1]^2 - 1,$$

and so an N-atom Morse cluster M_N has a potential energy given by

$$Morse_N(X;\rho) = \sum_{i<j} v_\rho(\|X_i - X_j\|).$$

Here X is the $3N$-dimensional vector of atomic coordinates, and X_i is the 3-dimensional vector of coordinates of the ith atom. The potential has a single parameter ρ that determines the width of the well (Figure 1), thus allowing it to capture aspects of the interactions for a wide range of materials. For example, the Lennard-Jones potential has the same curvature at the bottom of the well when $\rho = 6$; for this value of ρ the structural behavior of the two systems is very similar. Approximating the interactions between C_{60} molecules by a pair potential gives $\rho = 13.6$, whereas appropriate values for interactions between alkali metals are typically about $\rho = 3.1$.

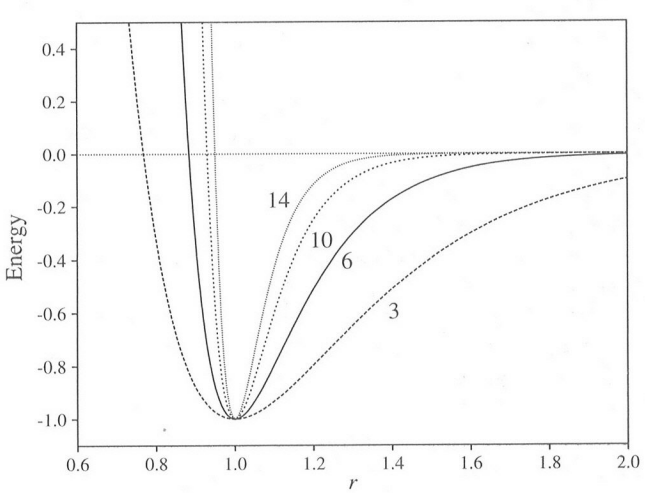

Figure 1 The Morse Potential for Different Values of ρ, as Labeled

The Morse clusters show a wide variety of structural behavior as a function of ρ. As ρ is increased the structures of the global minima change from polytetrahedral to Mackay icosahedral to decahedral to close-packed (Doye and Wales 1997). The potential-energy landscape also changes character with ρ. First, as ρ is increased the number of minima increases rapidly, changing the landscape from smooth to rough (Miller et al. 1999). On this basis alone, one expects the difficulty of global optimization to increase with ρ. Therefore, from the global-optimization perspective, $\rho < 6$ is the less interesting regime because optimization is easier than for Lennard-Jones clusters.

Second, the energy landscape is much more likely to have a multiple-funnel character at larger values of ρ (Doye et al. 2001b), further increasing the difficulty of global optimization. Usually, there are a variety of low-energy configurations, mainly, decahedral and close-packed, that are structurally very different and so well-separated in configuration space. This feature makes Morse clusters especially relevant to biomolecules, where, although there is an evolutionary pressure to make the energy landscapes of proteins sufficiently funnel-like that the protein can fold in a biologically appropriate time scale (Bryngelson et al. 1995), it is still common to find structures that act as traps hindering the folding process, both in experiment (Kiefhaber 1995, Radford et al. 1992) and simulations (Abkevich et al. 1994, Guo and Thirumalai 1995). Analyses of energy landscapes show that such traps correspond to additional funnels on the energy landscape (Levy and Becker 1998, Levy et al. 2001, Miller and Wales 1999, Mortenson et al. 2002).

The combination of these effects makes global optimization of Morse clusters at $\rho = 10$ and 14 a very challenging task, and one that is orders of magnitude more difficult than Lennard-Jones clusters. In the rest of this paper we will usually omit the word "putative." It is clear that except for trivial cases we have no proof of global optimality. A database of putative Morse cluster global minima is available up to $N = 80$ (Cambridge Cluster Database 2004). This was obtained through a combination of the application of global optimization algorithms and the generation of candidate structures on the basis of structural insight into the problem (Doye and Wales 1997). It has since been used as a global optimization test system (Roberts et al. 2000) but, until now, no unbiased global optimization algorithm has managed to obtain all the putative global minima. In this paper, we attempt just such a task.

2. Basin-Hopping and Funnel Landscapes

The landscape of a smooth potential function $E(X)$ with many local minima such as the Morse energy function can be viewed as a collection of basins of attraction, each associated with a local minimum. A gradient-based local optimization procedure such as a steepest descent, Newton, or quasi-Newton algorithm, when initiated at a point in a given basin of attraction, will converge to the associated local minimum. One strategy for the global minimization of such energy functions is to combine a standard gradient-based local search method for finding local minima with a global method for searching over the various basins of attraction. This general approach has come to be known as "basin-hopping" (Wales and Doye 1997).

Basin-hopping algorithms typically search the space of local minima by moving between adjacent local minima across a common basin boundary. This search strategy can be particularly effective on landscapes where the local minima are organized into a few large-scale funnels (Leary 2000). A funnel consists of a collection of local minima such that all monotonically descending sequences of successively adjacent local minima entering the funnel terminate at the funnel bottom. Note that all monotonically descending sequences must eventually enter a funnel, and clearly a downwardly biased search over adjacent local minima is a reasonable strategy to locate a funnel bottom. Note also that the global minimum lies in the set of funnel bottoms.

For any given point X and local search procedure LS, let $LS(X)$ be the local minimum of the energy function $E(X)$ that is obtained by applying LS to the starting point X. A typical basin-hopping algorithm produces a sequence of local minima from the iteration

$$X^{[i+1]} = LS(X^{[i]} + S),$$

where S is a random perturbation used to escape the current basin of attraction. The size of S is controlled to reach only nearby (and usually adjacent) local minima. Additionally, an acceptance/rejection step that depends on the energy change $E(X^{[i+1]}) - E(X^{[i]})$ is applied to bias the search toward lower energy values. Thus the sequence $\{X^{[i]}\}$ can be viewed as a Markov chain on a finite state space consisting of the local minima of $E(X)$. The transition probabilities are determined by the distribution of the perturbations S as well as the details of the acceptance/rejection rule.

In this paper we consider three forms of basin-hopping:

(i) BH is the original basin-hopping algorithm of Wales and Doye (1997). The same algorithm was developed independently in the protein-folding community and called *Monte Carlo with minimization* (Li and Scheraga 1987). The point $X^{[i+1]}$ is accepted or rejected according the Metropolis criterion, that is, all downhill steps ($E(X^{[i+1]}) < E(X^{[i]})$) are accepted, and uphill steps ($E(X^{[i+1]}) > E(X^{[i]})$) are accepted randomly with probability

$$p = e^{(E(X^{[i+1]}) - E(X^{[i]}))/T}.$$

Here, $T > 0$, is a preselected "temperature" parameter that remains fixed (i.e., no annealing is done) throughout a given run. If the step is rejected, new random perturbations S are generated until a step is accepted. Starting from a random initial local minimum, the Markov chain is extended for as long as desired (usually at least several thousand steps), and there are usually no absorbing states if the distribution of S is properly selected. A putative global minimum is the lowest energy state visited.

(ii) MSBH (monotonic sequence basin-hopping) is a variant due to Leary (2000) that attempts to exploit the funnel structure of the landscape more aggressively than BH. MSBH is basically BH at a temperature $T = 0$, i.e., only downhill steps are accepted, so the sequence of states is monotonically descending in energy. Once the sequence enters a given funnel, it will proceed on an energetically descending path to the funnel bottom, which is an absorbing state in the Markov chain. After a funnel bottom is recognized by the lack of progress in a sufficiently large number of attempted moves, the Markov chain is restarted from a new randomly selected local minimum. MSBH can be viewed as a multistart approach that samples funnel bottoms. It can be particularly effective if the number of funnel bottoms is small compared to the number of local minima, and the probability for entering the correct funnel associated with the global minimum is reasonably large.

(iii) TP-MSBH (two-phase MSBH), due to Locatelli and Schoen (2003), who introduce a modified energy function $E_G(X) = E(X) + G(X)$ where $G(X)$ is a geometrically-inspired penalty term. Let $LS_G(X)$ be the local minimum of $E_G(X)$ obtained by initiating a local search procedure at X. $LS_G(X)$ is then used as the starting point for a local search of $E(X)$, i.e., $X^{[i+1]} = LS(LS_G(X^{[i]} + S))$. Thus two local searches, one on the landscape of $E_G(X)$ and a second on the landscape of $E(X)$, are performed at each step (see §3 for more details). As in MSBH, the step is accepted if it improves the energy $E(X)$, and otherwise a new perturbation is attempted. The entire procedure is restarted from a new random local minimum of $E(X)$ after a prescribed number of step attempts with no progress. The effect of the first-phase local optimization on the landscape of $E_G(X)$ is to bias the search toward local minima of $E(X)$ with specific shape-related geometrical properties. As seen below, this can

greatly increase the probability of entering the proper funnel and finding the global minimum.

Wales and Doye (1997) note that basin-hopping algorithms for the global minimization of $E(X)$ in effect search a transformed piecewise constant "plateau" energy function $P(X) = E(LS(X))$ where the landscape of $P(X)$ consists of interlocking plateaus, each covering a given basin of attraction at a height equal to the energy of the local minimum in the basin. Much of the power of basin-hopping methods derives from the fact that, unlike global search methods applied directly to $E(X)$, there are no energetic barriers for moving downhill between plateaus, and basin crossing can easily take place at any point along a basin boundary with no necessity to search for special boundary-crossing points such as first-order transition points.

3. Two-Phase Local Searches

In this section two-phase (TP) local searches, introduced in §2, are described in more detail. The formal scheme of TP local searches is the following.

Procedure: TwoPhaseLocOpt(U);
Phase I: find

$$Y = [Y_1, \ldots, Y_N] = \arg \text{local} \min Morse_N^{mod}(X)$$

using $U = [U_1, \ldots, U_N] \in R^{3N}$ as the starting point;
Phase II: find

$$Z = [Z_1, \ldots, Z_N] = \arg \text{local} \min Morse_N(X; \rho) \quad (1)$$

using Y as a starting point;
End: return(Z).

Given a starting point U, the first phase returns a local minimum of a modified potential $Morse_N^{mod}$, while in the second phase such local minimum becomes the starting point of a local search driving to a local minimum of the original potential. While the second phase has the important task of driving the search back to the space of local minima of the original potential, the first phase can considerably increase the speed at which optimal clusters are detected, in particular through the choice of a modified potential that allows the funnel structure of the landscape to be modified in such a way that the funnel containing the global minimum is favored with respect to the others (Doye 2000). Therefore, the efficiency of TP local searches strongly depends on the modified potential. In this paper we employ the following modified potential (see Locatelli and Schoen 2002, 2003):

$$Morse_N^{mod}(X) = \min \sum_{i<j} (v_{\rho'}(\|X_i - X_j\|) \quad (2)$$

$$+ \mu \|X_i - X_j\| \quad (3)$$

$$+ (\max\{0, \|X_i - X_j\|_W^2 - D^2\})^2), \quad (4)$$

where

$$\|X_i - X_j\|_W = \left(\sum_{k=1}^{3} w_k (X_i^k - X_j^k)^2 \right)^{1/2}$$

is the weighted distance between two atoms. Note that ρ', μ, and D are nonnegative parameters, while $w_1 = 1$, $w_2, w_3 \in [0, 1]$, and $w_2 \geq w_3$. In the following, we comment on the different terms of the modified potential and the different impact that such terms have on the performance of TP-MSBH. Term (2) is the original Morse potential but with a parameter ρ' that is possibly different with respect to ρ. Parameter ρ' allows us to control the degree of smoothness of the modified potential: As ρ' increases the landscape of the modified potential becomes rougher and the number of local minima increases. The computations show that the choice $\rho' = \rho$ is often quite appropriate, although in some cases there may be some advantage from using $\rho' \neq \rho$.

Term (3) is a linear penalization of distance between atoms (typical values for μ range between 0 and 0.5). Its effect is that of disfavoring clusters containing atom pairs at large distance, thus favoring compactness of the clusters. While the use of this term gave some benefits during the optimization of Lennard-Jones clusters (see Locatelli and Schoen 2003), its relevance for the optimization of Morse clusters with large ρ values is less obvious and, as we will see, in the computations, setting $\mu = 0$ is often a feasible choice.

Term (4) is undoubtedly the most relevant one. It is a weighted-diameter penalization term, i.e., it strongly penalizes large weighted distance between atoms; more precisely, it penalizes weighted distances that are significantly larger than D (parameter D is often chosen between one or two units smaller than the expected diameter of the cluster, whose increase is proportional to $N^{1/3}$ and whose estimated value is returned by appropriate formulae). It also favors compact structures but whose shapes can be very different according to the choice of parameters w_2 and w_3. Indeed, if $w_2 = w_3 = 1$, then we are favoring spherical shapes, but as we decrease one or both the weights w_2 and w_3 we favor more oblate or prolate shapes. The advantages of using different weights w_2 and w_3 were already noticed in some cases for the optimization of Lennard-Jones clusters (Locatelli and Schoen 2003), but for the very difficult Morse clusters with large ρ values the advantages are much more evident because of the larger variability of the shapes of these clusters. An important question about parameters w_2 and w_3 is the following: Are there unique values for the weights w_2 and w_3 that are appropriate for *all* the Morse clusters? Unfortunately, the answer is no. As it will become clear in §4, weights w_2 and w_3 that are

extremely successful for some N values completely fail for other N values. The optimal choice of w_2 and w_3 for a given N value seems to be related to the moments of inertia of the corresponding optimal cluster (see §5) and these moments vary considerably with N for Morse clusters with large ρ value. This may seem to be bad news because we, of course, do not know optimal clusters in advance and, therefore, cannot choose the best values for w_2 and w_3. However, TP-MSBH appears to be quite a robust method with respect to the choice of w_2 and w_3. Indeed, even choices that are not too close to the moments of inertia often allow us to detect the optimal clusters very efficiently. In spite of the large variability of the moments of inertia, the computational results in §4 will show that almost all optimal Morse clusters at $\rho = 14$ for N up to 80 can be detected by TP-MSBH coupled with forward/backward procedures using only 5 different sets of parameters, often much more efficiently than with MSBH.

Note that the modifications to the original potential function introduced in the first phase step typically have a smoothing effect. However, the methodology and motivation for individual terms are quite distinct from standard smoothing methods (Morè and Wu 1996).

4. Computational Results

The first extensive computations of global minima of Morse clusters were described in Doye and Wales (1997) and made available via the Cambridge Cluster Database (2004). These include putative global minima for $N \leq 80$ and $\rho = 3$, 6, 10, and 14. The global minima were found by a combination of the BH algorithm and direct construction guided by the authors' experience with global minima of the Morse and other potentials. BH by itself was able to find all the initial results for $\rho = 3$, 6, and 10, and all but six for $\rho = 14$. Later improved global minima for $N = 61$ and 68, $\rho = 10$ (61B′ and 68D′ in the Cambridge Cluster Database) were found by W. Branz and noted in Doye et al. (2001b) and for $N = 30$ with $\rho = 14$ (30D′) by D. J. Wales using BH.

The MSBH algorithm was applied in the current study and found all the $\rho = 3$, 6, and 10 cases, including 61B′ and 68D′, and most of the $\rho = 14$ cases, including 30D′. MSBH also discovered the new 78D′, $\rho = 6$ "double icosahedron" putative global minimum. The $\rho = 3$ cases proved to be far easier than comparably-sized Lennard-Jones (LJ) clusters, while the $\rho = 6$ cases were of similar difficulty. However, particularly for $N > 40$ and $\rho = 10$ and 14, a few cases required extraordinary computational effort as measured by the number of trials (monotonic sequence starts from a random initial state) which is almost two orders of magnitude more than the most difficult Lennard-Jones clusters in the same size range. All of the Lennard-Jones optimal clusters were found with MSBH within 1,000 trials (Leary 2000), whereas quite a few of the Morse cases for the larger values of ρ required more than 1,000 trials for a single hit. In five instances for the larger clusters at $\rho = 14$, current putative Morse global minima have not yet been obtained by MSBH. Thus MSBH falls slightly short of meeting the challenge implicit in the observation in the CCDB that no unbiased global optimization method has been able to find all of the current putative global minima for $N \leq 80$ for the four ρ values. Here, unbiased means not incorporating detailed foreknowledge of the properties and morphological forms of known Morse global minima. Examples of biased algorithms include searching over highly restricted domains such as fcc, Mackay icosahedral, or Marks decahedral lattices (Northby 1987, Romero et al. 1999) which would necessarily preclude hitting any global minima outside those domains, and seeding algorithms with structures likely to be related to the global minima being sought (Wolf and Landman 1998).

In the remainder of this section we present and comment on computational results obtained using TP-MSBH, which show that that algorithm easily meets the Cambridge Cluster Database challenge. For economy of presentation, we restrict attention here primarily to $41 \leq N \leq 80$ and $\rho = 14$, the most difficult cases for MSBH. However, we have successively applied TP-MSBH to all the remaining cases covered in the Cambridge Cluster Database, as well as some intermediate values of ρ for which putative global minima structures are given in the Cambridge Cluster Database but explicit global optimizations have not previously been performed. The modified Morse function (see §3, expressions (2)–(4)) depends on a number of adjustable parameters, the most important of which are the diameter-penalty weights w_2, w_3 and the diameter threshold D. Here we fix $\rho' = \rho$ and $\mu = 0$, except when using the spherically symmetric weight set with all weights equal to 1, in which case we set $\mu = 0.1$ (the use of μ is not strictly necessary, but we have found the fixed setting $\mu = 0.1$ to be more useful than not in the spherically symmetric case). As discussed in §3, good ranges for D can be found systematically from N, so the diameter-penalty weights are the primary parameters of interest.

Our testing has shown that a significant range of diameter-penalty weights are usually effective for a given case, and a close match to an "optimal" set of weights is not necessary for successful application of the TP-MSBH algorithm. Thus a very general and effective algorithm can be formulated by simply spreading a sufficiently large number of weight sets fairly uniformly over the relevant portions of

Table 1 **The Five Sets of Parameters Employed in the TP-MSBH Computations**

	D	μ	w_2	w_3
P1	$3.5(N \leq 60)-4$	0.1	1	1
P2	$3(N \leq 60)-3.5$	0	1	0.7
P3	3	0	0.8	0.8
P4	2.5	0	0.75	0.5
P5	$2(N \leq 60)-2.5$	0	0.5	0.5

Note. The dependence of the parameter D on cluster size is given in parentheses.

the weight space so that at least one set is likely to be effective. Because the weight space is of low dimension (in effect, two-dimensional because w_1 is fixed at 1.0) and the relevant region $0.0 \leq w_3 \leq w_2 \leq 1.0$ quite small, the algorithm can effectively sample weight space with a manageable number of weight sets.

After a very limited amount of experimentation, we found that the 5 sets of parameters P1–P5 given in Table 1 were sufficient to find all but one ($N = 61$) of the cases of interest using 50 trials for each of the parameter sets. The results are reported in Table 2, where each column P1–P5 reports the number of successes in 50 random runs of TP-MSBH with the given set of parameters. We do not report CPU times, which change too rapidly with technology, and prefer to use the number of two-phase local searches as the measure of computational work. For purposes of comparison with other algorithms that use repeated local optimizations, a two-phase search is only slightly more costly (around 25%) than a single local optimization of the unmodified Morse potential, and a single run of TP-MSBH at the cluster-size range here corresponds to at most 1,000 two-phase searches (and often far fewer). We can add that in the size range tested (between 50 and 80 atoms), each local optimization consists of a minimization of a smooth function of between 150 and 240 variables, which typically requires less than one CPU second on a standard PC.

Table 2 **Number of Successes in 50 Random Runs for the Five Different Parameter Sets**

N	P1	P2	P3	P4	P5	FB	N	P1	P2	P3	P4	P5	FB
41	0	0	0	2	0		61	0	0	0	0	0	
42	0	0	0	0	0	*f*	62	0	4	0	0	0	
43	0	0	0	1	0		63	0	21	0	25	0	
44	0	0	0	0	0	*f*	64	1	18	0	26	0	
45	0	0	0	0	0	*b*	65	0	6	0	3	0	
46	0	0	0	0	3		66	0	8	0	0	0	
47	0	0	0	0	1		67	0	3	0	0	0	
48	0	0	0	4	28		68	5	0	0	0	0	
49	0	0	0	29	14		69	0	1	8	1	0	
50	0	47	0	5	0		70	0	7	5	7	0	
51	0	31	0	13	0		71	3	2	49	12	0	
52	0	21	1	4	0		72	0	0	45	7	0	
53	1	4	1	0	0		73	1	0	25	1	0	
54	0	0	0	0	0	*b*	74	2	0	47	18	36	
55	0	0	0	0	0	*b*	75	11	0	44	8	5	
56	1	0	0	0	0		76	2	0	38	3	6	
57	1	0	0	0	0		77	0	0	11	10	11	
58	39	1	12	0	0		78	0	0	0	5	4	
59	39	0	19	0	0		79	1	0	0	0	0	
60	25	0	20	0	0		80	0	0	0	0	0	*f*

For the sizes where the optimal cluster has not been detected with any of the 5 sets of parameters, column FB reports f if the optimal cluster has been detected by a forward procedure, and b if it has been detected by backward procedure from a putative global minimum found at an adjacent size with the same parameter set. The forward/backward procedures are commonly used in cluster-optimization studies to take advantage of the fact that the global optima for adjacent size N and $N+1$ sometimes (but by no means always) have very similar structures, differing only in the placement of usually one and at most a few atoms, typically on the surface. Thus the removal or addition of an atom at one size, in conjunction with some simple searches, is a useful technique for locating global minima at the adjacent sizes. Here we limit the procedure to 100 random additions/deletions of a single atom to/from the best configuration with $N-1/N+1$ atoms found by TP-MSBH. For the sake of completeness we point out that the optimal configurations found by the forward/backward procedures (those for $N = 42$, 44, 45, 54, 55, and 80) can also be found directly by TP-MSBH by simply increasing the number of random trials (never more than 1,000 and usually far fewer).

Along with the "obvious" choice of the spherically symmetric weight set $w_1 = w_2 = w_3 = 1.0$ in P1, the remaining four weight sets in P2–P5 were selected primarily on the basis of broad coverage of the weight space (see Figure 2) rather than by tuning for specific results. The five sets of parameters employed here are not unique or the end product of an extensive parameter selection and optimization process, and certainly others with good coverage of the weight-space will work as well. Also, filling in gaps in the weight space coverage of P1–P5 will give improved performance. For example, the single missed case at $N = 61$ is easily found in two of 50 trials using a weight set P6 midway between P3 and P5. Also, the choice $w_2 = 0.6$, $w_3 = 0.4$, and $D = 2$ in an area of weight space not well covered by P1–P5 allows us to detect all optimal clusters for $N = 41$–47 with much higher success rates and without the use of forward/backward procedures. However, we emphasize that our aim is not to find optimal parameters for each N, but to show that a very limited number of parameter sets is sufficient to find all optimal clusters and that an efficient and robust algorithm can be constructed without

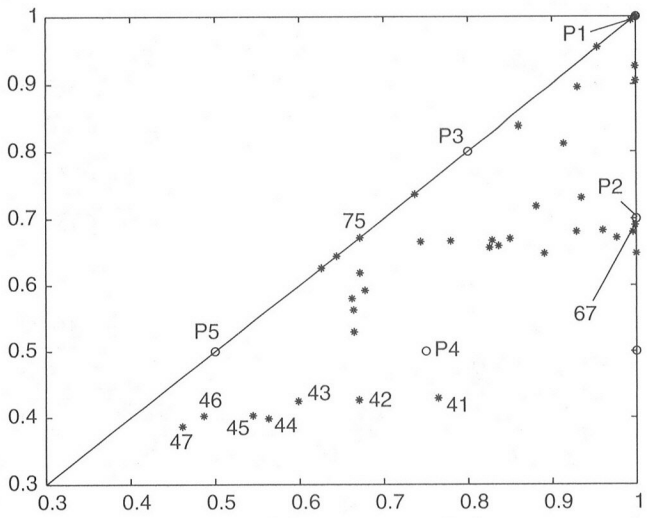

Figure 2 **Moment-of-Inertia Weights w_2 (Horizontal Axis) and w_3 (Vertical Axis) for $N = 41$–80, $\rho = 14$.**

Note. They are plotted as * with coordinates (w_2, w_3). Weight sets P1–P5 used to find all global minima are plotted as ∘. Cases discussed in the text are labeled with the corresponding N value.

excessive parameter tuning or foreknowledge of the global minima being sought.

We have confirmed the effectiveness, efficiency, and robustness of the TP-MSBH approach proposed here by easily finding all previously known optimal clusters in the Cambridge Cluster Database. Further evidence of these properties is provided by the discovery or rediscovery by TS-MSBH of several new putatively optimal clusters with improved energies relative to those known prior to the current study. For $N = 78$ at $\rho = 6$ a new optimal cluster was first discovered by MSBH, and later also found with a much higher success rate by TP-MSBH. For $N = 24$, 25, 45, 48, and 51 at $\rho = 8$, new optimal clusters have been discovered by TP-MSBH. All these new clusters are shown and described in §6.

Several of the cases $N = 41$–80 and $\rho = 14$ discussed in this section have proven to be enormously difficult for the BH and MSBH algorithms. For more than half of the cases where $N > 60$, 1,000 trials of the MSBH algorithm have proven to be insufficient, and several required more than 10,000 trials to achieve a single success. In five cases, neither version of single-phase basin-hopping has to date been successful despite extensive computational effort. In contrast, the TP-MSBH algorithm has consistently found all the global minima with only a few hundred trials at each N value spread over a very limited number of parameter sets selected for broad coverage of the parameter space. Thus, the power and generality of the two-phase idea with a weighted-diameter penalization is demonstrated by the reliably large amplification relative to single-phase approaches of the probability of hitting the global minimum with at least one parameter set. This far outweighs the dilutive effect of distributing a given number of trials over multiple parameter sets.

5. Shape Factors and Diameter-Penalty Weights

The shape of a cluster can be approximately described by shape factors representing the ratios of the lengths of the minor, intermediate, and major axes of an ellipsoid with the same principal moments of inertia. The three-by-three moment-of-inertia tensor (also known as the *shape tensor* or *radius-of-gyration tensor*) for any configuration X of N identical atoms is defined as

$$Q_{mn} = \frac{1}{N}\sum_{j>i}(X_i^m - X_j^m)(X_i^n - X_j^n).$$

The eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \lambda_3$ of Q correspond to the squared lengths of the equivalent ellipsoid axes, with the eigenvectors representing the directions of the axes. The trace $\mathrm{Tr}(Q) = \lambda_1 + \lambda_2 + \lambda_3$ is the squared radius of gyration, i.e., the mean square distance of the atoms from the center of mass.

The weighted-diameter penalization term for any pair (i, j) of atoms is an increasing function of the quadratic form

$$\|X_i - X_j\|_W^2 = \sum_{k=1}^{3} w_k\left(X_i^k - X_j^k\right)^2$$

defined by the weights $1 = w_1 \ge w_2 \ge w_3 \ge 0$. The corresponding ellipsoids of constant weighted-diameter penalty are centered at the center of mass of the atom pair with principal axes of length proportional to $1/(w_k)^{1/2}$ and directions along the coordinate axes.

Suppose the eigenvalues of the moment-of-inertia tensor of the global minimum are known in advance, and we wish to find diameter-penalty weights that will work well in the TP-MSBH algorithm to find that global minimum. The obvious choice $w_1 = 1$, $w_2 = \lambda_1/\lambda_2$, and $w_3 = \lambda_1/\lambda_3$ matches the shape of the ellipsoids of constant weighted-diameter penalty to that of the moment-of-inertia ellipsoid, thus penalizing configurations with the "wrong" shape. Examples of moment-of-inertia weights (w_2, w_3) are plotted in Figure 2 for the cases $N = 41$–80, $\rho = 14$, along with the weight sets P1–P5 described in §4. The point (1, 1) corresponds to spherical shapes, the bounding line $w_2 = w_3$ to oblate shapes, and the bounding right edge $w_2 = 1$ to prolate shapes.

Empirically we have found that moment-of-inertia penalty weights are always a reasonable choice in the sense that all global minima have been found within 1,000 runs (and usually far fewer) with their use. Moreover, they are often very good or even near-optimal in terms of approximately maximizing

the probability of hitting the global minimum, and in difficult cases, the success of TP-MSBH depends on a reasonable match of the diameter-penalty weights to the optimal cluster shape. Conversely, a relatively few weight sets such as P1–P5 that, as seen in Figure 2, are fairly uniformly distributed over the space of moment-of-inertia shape factors, are sufficient to find all the global minima.

An example of the efficacy of matching penalty weights to moment-of-inertia shape factors is the new double icosahedral global minimum for $N = 78$ and $\rho = 6$ that was first found by MSBH. This case is in fact quite difficult for that algorithm, which hit it only three times in 1,000 runs. Using the moment-of-inertia weights $w_1 = 1.0$, $w_2 = 0.9718$, and $w_3 = 0.4957$ corresponding to the rather extreme prolate shape of this structure, the TP-MSBH algorithm hit it in 410 of 1,000 runs. Weight sets that are a poor match were far less successful. Similarly, exploration of the weight space for the relatively difficult $N = 67$, $\rho = 14$ case found the best weights attempted to be $w_1 = 1.0$, $w_2 = 1.0$, and $w_3 = 0.7$, in excellent agreement with the moment-of-inertia weights $w_1 = 1.0$, $w_2 = 0.9986$, and $w_3 = 0.6907$ for this prolate shape. Also, the oblate weight set P3 is the most successful of the five sets discussed in §4 for the oblate Marks decahedron at $N = 75$ and $\rho = 14$. Finally, extremely aspherical structures such as the $N = 41$–47 and $\rho = 14$ cases where the global minima have very flat decahedral shapes can be among the most difficult for TP-MSBH using just the parameter sets P1–P5. The only successful weight sets P4 and P5 for these cases are the best matches to the moment-of-inertia weights. However, as mentioned in the previous section, considerably improved results can be obtained with the addition of a weight set closer to the shape factors for theses cases.

6. New Global Minima

The new global minima that were discovered during this research program are depicted in Figure 3. All these new global minima are for $\rho = 8$ except M_{78}, which is for $\rho = 6$, and have been added to the Cambridge Cluster Database.

The 24-atom structure is formed from the 23-atom triple icosahedron by the addition of a capping atom at the edge of the cluster. The $\rho = 6$ global minimum instead has the capping atom at the other alternative four-fold coordinate site. In the 23-atom triple icosahedron, mentioned above, the centers of the three interpenetrating icosahedra form an equilateral triangle. The new M_{25} global minimum also consists of three interpenetrating icosahedra, but with their centres in an open "V"-shape. This structure has previously been found for clusters interacting with the Dzugutov potential (Doye et al. 2001a).

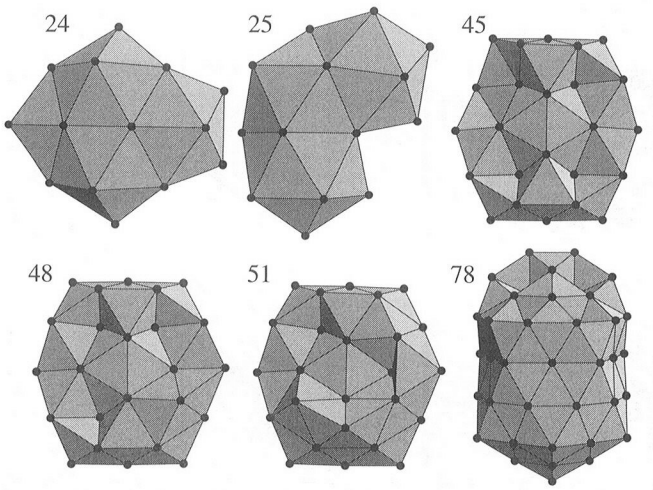

Figure 3 **New Morse Cluster Global Minima Found During this Research**

The 45-, 48-, and 51-atom clusters are all incomplete Mackay icosahedra, but which have a slightly different choice of vacant surface sites compared to the icosahedral global minima previously located at these sizes. In particular, all three structures can be formed from the icosahedral global minimum for the preceding size by the addition of an atom to a vertex site that is only five-coordinate, rather than to the edge site that completes the five-fold ring around that vertex.

At $N = 87$ a structure can be formed that is made up of two interpenetrating Mackay icosahedra and is related to the 19-atom double icosahedron in the same way that the 55-atom Mackay icosahedron is related to the 13-atom icosahedron. The new M_{78} global minimum is formed from this double icosahedron by the removal of two triangular faces.

7. Conclusion and Future Research

We have shown that the two-phase basin-hopping algorithm with shape-directing weighted diameter penalties is a highly effective and robust global optimization methodology for Morse clusters. A very limited parameter grid selected without detailed parameter tuning is sufficient to detect all known optimal clusters and also to find new ones.

However, we have also shown that there seems to exist a "natural" choice of the weight parameters based on the cluster-shape factors determined by the principal moments of inertia. One possible aim for future research is to investigate more sophisticated strategies than the current grid search that perhaps incorporate adaptive elements exploiting this connection. For example, the current weighted-diameter penalty term is applied in a space-fixed, rather than molecule-fixed, coordinate system. Thus the corresponding compressional forces will be changed in the

molecule-fixed system by simply rotating the cluster in the space-fixed system, which does not seem appropriate if indeed we want to search over the space of cluster shapes. The insertion of an explicit rotation step after each local optimization to align the principal moment-of-inertia axes of the current iterate with the coordinate axes would more closely couple the diameter penalty to cluster shape during the course of the optimization.

Although the transformations of the Morse potential outlined here have been designed with clusters in mind, competing conformations in biomolecules are also likely to differ in their shape. Therefore, these methods are potentially transferable to the global optimization of biomolecules, such as in the protein-folding problem. Even if the particularly penalty functions that we use here are not the most appropriate, the principles of our approach are still likely to apply, namely, the design of potential modifications that can discriminate between competing structures, which are then used in a series of optimization runs that favor different regions of shape space. Indeed, preliminary applications to protein-protein docking (Addis and Schoen 2003) have been encouraging.

Acknowledgments

Jonathan P. K. Doye is grateful to the Royal Society for financial support. Robert H. Leary acknowledges support from the National Science Foundation under Cooperative Agreement ACI9619020, and also the contributions of Kevin Chan of Harvard College who performed many of the MSBH runs, including those that found the new structure at $N = 78$, $\rho = 6$. Marco Locatelli and Fabio Schoen acknowledge support from Progetto FIRB "Ottimizzazione Non Lineare su Larga Scala" and also the contribution of B. Addis, Università di Firenze, for her support and suggestions.

References

Abkevich, V. I., A. M. Gutin, E. I. Shakhnovich. 1994. Free energy landscape for protein folding kinetics: Intermediates, traps and multiple pathways in theory and lattice model simulations. *J. Chem. Phys.* **101** 6052–6062.

Addis, B., F. Schoen. 2003. A randomized global optimization method for protein-protein docking. Technical report DSI 4-2003, Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Firenze, Italy.

Bryngelson J. D., J. N. Onuchic, N. D. Socci, P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* **21** 167–195.

Doye, J. P. K. 2000. The effect of compression on the global optimization of atomic clusters. *Phys. Rev. E* **62** 8753–8761.

Doye, J. P. K., D. J. Wales. 1997. Structural consequences of the range of the interatomic potential: A menagerie of clusters. *J. Chem. Soc. Faraday Trans.* **93** 4233–4244.

Doye, J. P. K., D. J. Wales. 1998. Thermodynamics of global optimization. *Phys. Rev. Lett.* **80** 1357–1360.

Doye, J. P. K., M. A. Miller, D. J. Wales. 1999. Evolution of the potential energy surface with size for Lennard-Jones clusters. *J. Chem. Phys.* **111**(18) 8417–8428.

Doye, J. P. K., D. J. Wales, S. I. Simdyankin. 2001a. Global optimization and the energy landscapes of Dzugutov clusters. *Faraday Discuss* **118** 159–170.

Doye, J. P. K., D. J. Wales, W. Branz, F. Calvo. 2001b. Modelling the structure of C_{60} clusters. *Phys. Rev. B* **64** 235409, 1–11.

Guo, Z. Y., D. Thirumalai. 1995. Kinetics of protein-folding—Nucleation mechanism, time scales, and pathways. *Biopolymers* **36** 83–102.

Kiefhaber, T. 1995. Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci. USA* **92** 9029–9033.

Lau, K. F., K. A. Dill. 1989. A lattice statistical-mechanics model of the conformational and sequence-spaces of proteins. *Macromolecules* **22** 3986–3997.

Leary, R. H. 2000. Global optimization on funneling landscapes. *J. Global Optim.* **18** 367–383.

Levy, Y., O. M. Becker. 1998. Effect of conformational constraints on the topography of complex potential energy surfaces. *Phys. Rev. Lett.* **81** 1126–1129.

Levy, Y., J. Jortner, O. M. Becker. 2001. Dynamics of hierarchical kinetics on the energy landscapes of hexapeptides. *J. Chem. Phys.* **115** 10533–10547.

Li, Z., H. A. Scheraga. 1987. Monte-carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. USA* **84** 6611–6615.

Locatelli, M., F. Schoen. 2002. Fast global optimization of difficult Lennard-Jones clusters. *Comput. Optim. Appl.* **21** 55–70.

Locatelli, M., F. Schoen. 2003. Efficient algorithms for large scale global optimization: Lennard-Jones clusters. *Comput. Optim. Appl.* **26** 173–190.

Miller, M. A., D. J. Wales. 1999. Energy landscape of a model protein. *J. Chem. Phys.* **111** 6610–6616.

Miller, M. A., J. P. K. Doye, D. J. Wales. 1999. Structural relaxation in Morse clusters: Energy landscapes. *J. Chem. Phys.* **110** 328–334.

Morè, J. J., Z. Wu. 1996. Smoothing techniques for macromolecular global optimization. Giannessi, Di Pillo, eds. *Nonlinear Optimization and Applications*. Plenum Press, New York, 297–312.

Morse, P. M. 1929. Diatomic molecules according to the wave mechanics, II. Vibrational levels. *Phys. Rev.* **34** 57–64.

Mortenson, P. N., D. J. Wales, D. A. Evans. 2002. Energy landscapes of model polyalanines. *J. Chem. Phys.* **117** 1363–1376.

Northby, J. A. 1987. Structure and bonding of Lennard-Jones clusters: $13 \leq n \leq 147$. *J. Chem. Phys.* **87** 6166–6177.

Radford, S. E., C. M. Dobson, P. A. Evans. 1992. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358** 302–307.

Roberts, C., R. L. Johnston, N. T. Wilson. 2000. A genetic algorithm for the structural optimization of Morse clusters. *Theoret. Chemistry Accounts* **104** 123–130.

Romero, D., C. Barrón, S. Gómez. 1999. The optimal geometry of Lennard-Jones clusters: 148-309. *Comp. Phys. Comm.* **123** 87–96.

The Cambridge Cluster Database. 2004. http://www-wales.ch.cam.ac.uk/CCD.html.

Wales, D. J., J. P. K. Doye. 1997. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **101** 5111–5116.

Wales, D. J., H. A. Scheraga. 1999. Global optimization of clusters, crystals and biomolecules. *Science* **285** 1368–1372.

Wolf, M. D., U. Landman. 1998. Genetic algorithms for structural cluster optimization. *J. Phys. Chem. A* **102** 6129–6137.