

Minimization of the potential energy surface of Lennard–Jones clusters by quantum optimization

Thomas Gregor^{a,b,*}, Roberto Car^b

^a Princeton University, Lewis-Sieglar Institute for Integrative Genomics, Carl Icahn Laboratory, Princeton, NJ 08544, USA

^b Department of Chemistry, Princeton Institute for the Science and Technology of Materials (PRISM), Princeton University, Princeton, NJ 08544, USA

Received 28 April 2005; in final form 15 June 2005

Available online 18 July 2005

Abstract

One of the most widely used strategies for global optimization employs the concept of classical simulated annealing. In the last decade an alternative approach has been suggested based on quantum simulated annealing. Here, we apply quantum annealing ideas to finding minimum energy structures of Lennard–Jones clusters. We find that quantum annealing is superior to classical simulated annealing but is affected by ergodicity breaking difficulties similar to classical simulated annealing. This difficulty is particularly serious for larger clusters with multiple funnel potential energy surfaces.

© 2005 Elsevier B.V. All rights reserved.

Finding the atomic structure that corresponds to the global minimum of the potential energy surface (PES) is a central problem in cluster physics. It is highly non-trivial as for most global optimization problems: (i) the PES has many local minima, whose number grows exponentially with cluster size and (ii) the PES usually has a multi-funnel structure, reflecting the simultaneous presence of competing growth sequences. Both features are present in Lennard–Jones (LJ) clusters, in spite of the relative simplicity of the inter-atomic interactions given by the two-body potential

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

where ϵ and σ are reduced LJ units and r_{ij} is the Euclidean distance between atoms i and j . The many barriers that separate local minima in the PES of the cluster are a result of the repulsive hard core which severely

hampers local atomic rearrangements. At the same time, at least two main funnels have been identified in the PES of relatively small (up to few hundred atoms) LJ clusters: they reflect a growth sequence which, in this size range, is dominated by competing icosahedral and decahedral or cubic close-packed (ccp) structural motifs [1,2].

Different optimization methods have been tested and applied to LJ clusters. These optimization methods can be divided into two classes based on the criteria that: (1) they use the PES as the only guiding function (objective function) or (2) use additional information such as knowledge about the dominant growth sequences. Among the methods belonging to the first class classical simulated annealing (SA) or its variants [3] are widely used. SA relies on thermal fluctuations to overcome barriers and achieve global sampling of the PES. More recently, however, it has been recognized that quantum fluctuations can be superior in some cases to classical thermal fluctuations in order to achieve global optimization. The corresponding approaches have been called quantum annealing (QA) methods [4–8].

* Corresponding author.

E-mail addresses: gregor@princeton.edu (T. Gregor), rcar@princeton.edu (R. Car).

Interestingly, what is perhaps the very first proposal for QA was made in the context of LJ clusters [9]. This pioneering work was limited to very small clusters (up to 19 atoms) and no comparison was made with alternative strategies based on classical SA. In this work, we provide a quantitative comparison of QA and SA methods for LJ clusters. In particular, we assess how alternative methods perform vs. system size and how they handle multi-funnel PESs. In agreement with similar findings for random ferromagnets [4], random Ising models [5,7], the traveling salesman problem [8] and simple protein models [6], we find QA to be superior to SA in the LJ cluster problem. However, both QA and SA protocols become unable to find the global minimum at a reasonable cost when the size of the system becomes too large, irrespective of the single- or multi-funnel nature of the PES. This is fully consistent with theories of annealing methods according to which exponentially large annealing times would be required to find the global minimum of a rugged PES of a large many-body system, regardless of its complexity. As an additional result of our analysis we propose a new optimization scheme that we call the Replica Pinned Quantum Annealing (RPQA) method which, for LJ clusters, is superior, albeit only quantitatively, to all the other schemes investigated in the present work. Using RPQA we were able to find the global minima¹ of icosahedral clusters with up to 201 atoms, which is well beyond the capability of common SA and QA approaches.

In SA one generates atomic configurations distributed as $\exp(-\sum_{ij}\Phi_{ij}/T)$ using Monte Carlo (MC) or Molecular Dynamics. The temperature T is a parameter to control the size of thermal fluctuations. Starting at high T , where the system explores liquid-like equilibrium configurations, the temperature is slowly reduced until the system gets trapped into a local minimum of low energy. QA is conceptually very similar. In the LJ case, one constructs a quantum Hamiltonian H by adding to the PES a kinetic energy \hat{K} given by:

$$\hat{K} = \sum_i -\frac{\hbar^2}{2m} \nabla_i^2, \quad (2)$$

where ∇_i^2 is the Laplacian operator acting on the i th atom. The quantum coupling $\gamma = \hbar^2/2m$ is used as a parameter to control the size of quantum fluctuations. Simulations are performed at zero or at small finite T using a quantum MC approach. At large values of γ quantum fluctuations delocalize the system over a large portion of configuration space. Subsequently, γ is slowly reduced until the system gets trapped into a local minimum configuration. We use standard Metropolis MC to perform classical SA and a path integral Monte Carlo

(PIMC) approach to perform QA. In PIMC, one exploits the isomorphism between a quantum system at finite T and a classical ring polymer consisting of P replica of the original system at the same temperature [11–14]. Each replica has a reduced potential energy Φ/P and nearest neighboring replica are coupled harmonically with a force constant $PT^2/4\gamma$, resulting from a discretized Feynman path integral. A larger value of P is required for a system with a large quantum character, i.e., a system with large γ and/or small T . In our PIMC implementation, we follow the bisection procedure of [15] to sample polymer configurations but alternative procedures based on staging can be used as well [16]. We take a value of $T = 1/\epsilon$ in reduced LJ units. This value of T is sufficient to discriminate between the lowest-lying local minima of the PES in the entire size range considered in the present investigation. Furthermore, it allows us to use a value of $P = 100$ in most of our simulations where \hbar is varied between 0 and 12 atomic units (a.u.). Because different replica are coupled only when they are nearest neighbors, the computational cost of a PIMC simulation is linear in P , thus a PIMC simulation is roughly P times more expensive than a classical MC simulation that explores the same number of configurations. In the following we take this into account when comparing classical and quantum annealing protocols. In all simulations, the starting configurations were taken from well-equilibrated configurations generated by classical MC at $T = \epsilon$. P equal replica were constructed from one of these configurations in the QA case.

Initially, we adopted a simple annealing schedule in which the control parameter (either T or γ) was reduced linearly with simulation time. While this procedure works well for small clusters, when the number of atoms is greater than 20 it becomes increasingly more difficult to avoid trapping in a local minimum. We rationalize these findings as follows: at sufficiently large $T(\gamma)$, the system ergodically samples the PES, but when $T(\gamma)$ becomes smaller than some threshold, ergodicity breaks, evolution becomes sluggish, and eventually the system becomes trapped in a local minimum. This is precisely what one expects for classical SA when many barriers must be crossed in order to reach the global minimum. A similar phenomenology appears in QA. In this case, it is a manifestation of the inability of the system to adiabatically follow the ground-state [7]. It turns out that a significantly more efficient strategy rests in letting the system evolve above the threshold for ergodicity breaking while at the same time systematically performing local optimizations from the configurations visited during ergodic evolution. We find that a good choice for ergodic evolution corresponds to $T = 0.18\epsilon^2$ in the SA

¹ It can never be proved that one has found the global minimum for a particular cluster. Here, we qualify as ‘global minima’ minima that correspond to the Cambridge Cluster Database [10].

² Recall that the melting temperature for small clusters is around $T = 0.2\epsilon$.

Table 1
Comparison of computational cost for different global optimization algorithms for LJ₃₈

Method	Cost (steps)
Classical annealing	>100000
Quantum annealing	>100000
Classical sampling	22732
Quantum sampling	3500

The numbers represent the average for 25 different initial conditions. ‘Annealing’ refers to a slow decrease in $T(\gamma)$, whereas in the ‘sampling’ protocol $T(\gamma)$ is kept constant and local minima are mapped from local optimizations by a conjugate gradient algorithm every 100 MC (PIMC) steps. Both classical and quantum annealing were unable to find the global minimum (> signifies that we stopped the simulation after that amount of steps).

case and to $\hbar = 12$ in the QA case. In both cases we map local minimum configurations from local optimizations by a conjugate gradient algorithm every 100 MC (PIMC) steps. A step corresponds to a single MC update of the coordinates of all the particles in a cluster in the classical case and of all the replicaes in the quantum case. This sampling strategy is compared to SA and QA with a linear annealing schedule for the special case of a 38 atom cluster in Table 1. In the sampling cases, the additional cost of performing the conjugate gradient optimizations has been taken into account in assigning a cost. The sampling approach is more efficient than standard annealing procedures. Similar results also apply to other cluster sizes and we adopt the sampling strategy in all subsequent simulations.

The performance of SA and QA (with the sampling strategy) is evaluated in Table 2 for selected cluster sizes. QA works significantly better than SA, particularly for large cluster sizes, where the savings can be beyond an order of magnitude. The advantage, however, is merely quantitative because for clusters larger than about 100 atoms, neither SA nor QA is able to find the global minimum. While the 38 and 75 atom clusters have ccp and decahedral global minima, respectively, all other clusters in the table have icosahedral global minima. Both LJ₃₈ and LJ₇₅ are examples of clusters, where the low energy PES is dominated by two funnels. Since icosahedral structures are significantly more numerous and easily accessible when T or γ are above the ergodicity threshold, these clusters are considerably more difficult to optimize, as recognized in [17]. The lowest icosahedral minimum which lies only slightly above the global ccp and decahedral minimum, respectively, is found most of the time. This difficulty with a multi-funnel PES is evident in Table 2, from which one sees that both SA and QA are able to find the global minimum of LJ₃₈ but not of LJ₇₅. However, even for ‘easier’ icosahedral clusters, both methods are unable to find the global minima of clusters of more than 100 atoms. We assign this difficulty to the exponential growth of the number of local minima and of the barriers connecting adjacent

minima with cluster size. This inevitably leads to ergodicity breaking in SA and QA optimization algorithms for a sufficiently large cluster size. The effect of a multiple funnel PES is that it leads to ergodicity breaking at smaller cluster sizes.

One key aspect of annealing methods that is being lost when using the sampling strategy is the iterative improvement. At high $T(\gamma)$, the system samples larger regions of configuration space that are successively restricted to smaller regions comprising lower lying minima as $T(\gamma)$ is lowered. In other words, the system improves the knowledge of the PES during its annealing history. Unfortunately, we have seen that this strategy ceases to be effective when the annealing time is finite and the number of local minima is very large. A simple way of incorporating learning from sampling history is achieved by employing RPQA. In this approach the system samples configuration space at $\hbar = 10$ a.u. while successive local optimizations are performed, as in the sampling method described above. However, this time one of the P replicaes is held fixed at the coordinates of the best local minimum found during the previous annealing history. As the sampling proceeds, a better local minimum is eventually found. Then the coordinates of the replica that happened to be in the catchment area of this minimum are pinned at their corresponding local minimum value while the replica that was held fixed at a higher-lying local minimum is released. By repeating this procedure many times, one of the replicaes will find the global minimum. We see in Table 2 that this procedure represents a significant improvement over the pure sampling procedure. By using the RPQA method the global minimum of a LJ₁₄₇ cluster can be found in 1 day using a standard laptop, a result that is well beyond the capability of the SA and QA sampling methods. Further improvements may be possible by using more elaborate pinning strategies (perhaps involving more than one replica at a time). Features of RPQA are reminiscent of the simulated tempering method which represents a significant improvement over simple classical simulated annealing strategies [18].

To explain why QA methods perform better than their classical counterpart, we analyze the so-called ‘inherent PES’ sampled by classical and quantum schemes. The concept of inherent PES was introduced in classical liquids to eliminate thermal effects from the PES [19]. It consists of a collection of the local PES minima that underlie liquid-like configurations. In Fig. 1, we reported the distribution of low-lying local minima found by classical and quantum sampling methods for LJ₁₄₇. As we can see, the quantum inherent PES has more weight on low energy states, exhibiting a tail extending to lower energy configurations. This is at first surprising because the corresponding configurations before local optimization have systematically higher energies in the quantum case. However,

Table 2

Comparison of the computational cost of classical, quantum and RPQA search methods for different cluster sizes

Size	Classical sampling	Quantum sampling	RPQA
LJ ₂₆	359	168	144
LJ ₃₈	22 732	3500	3100
LJ ₅₅	14 387	1800	1400
LJ ₇₀	>500 000	88 900	30 800
LJ ₇₅ ^a	>1 000 000	37 900	34 200
LJ ₁₀₇	>1 000 000	>1 000 000	251 100
LJ ₁₄₇	>1 000 000	>1 000 000	229 200
LJ ₁₆₇	>1 000 000	>1 000 000	618 000
LJ ₂₀₁	>1 000 000	>1 000 000	899 200

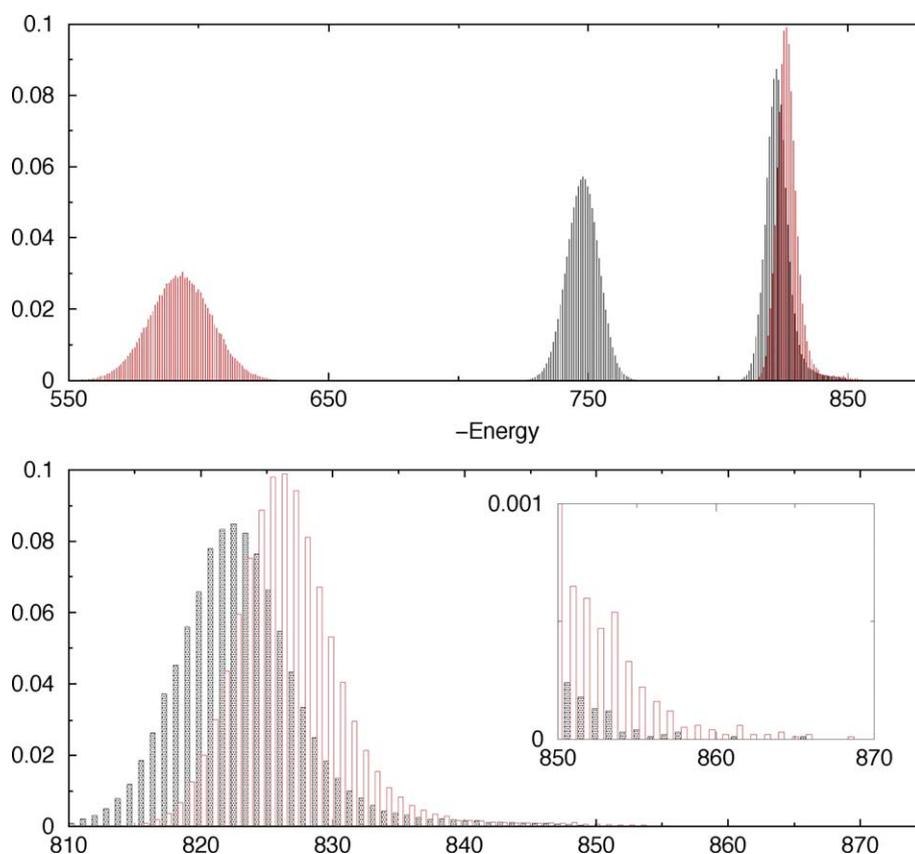
^a Values are for lowest icosahedral minimum.

Fig. 1. Energy distribution of 100 replica for a LJ₁₄₇ cluster for classical (black) and quantum (red) sampling. Notice the inverted energy scales (more stable structures are on the right). Upper panel: Energy distribution before (left peak) and after (right peak) local minimization. Lower panel: Inherent PES (larger view of right most peaks from upper panel). Inset: Enlargement of lowest energy configurations. The global minimum would be at -876ϵ . (For interpretation of the references to colour in this figure legend, the reader is referred to the online version of this article.)

the main reason for this higher energy is that quantum configurations include classically forbidden configurations of repulsive potential energy. These are associated to barrier crossings which are made easier by quantum fluctuations. For this reason, quantum sampling has at the same time a consistently higher weight than classical sampling on low-lying minima that are separated by high and narrow barriers from adjacent regions of configuration space. We found that these barriers are usually associated with small local readjustments leading to elimination of defects. These local processes become

increasingly more difficult at large cluster size when many routes to disorder are present and eventually the system is trapped in a glassy configuration. RPQA obviates this difficulty to some extent. In all cases, we find that at sufficiently low energy, the replica that are pinned successively are just two adjacent replica, each one of which in turn helps the other to find a better minimum. This cooperative effect between adjacent replica allows effective defect annealing and corresponding downhill motion until the global minimum is found.

Finding the global minimum is considerably more difficult when the system has to move from the icosahedral to the ccp funnel at low energy. As we can see from Table 2 none of the methods considered here succeeded in achieving this result within a reasonable time for LJ₇₅, while all the methods worked successfully for LJ₃₈. The reduced success rate for LJ₇₅, as compared to LJ₃₈, is partially due to the large increase in the number of local minima with cluster size. However, a less trivial reason is associated to the presence of two funnels in the PES. To better investigate this issue, we performed quantum sampling runs with $P = 500$ and $\hbar = 10$, where one of the replica was pinned at the configuration of either the global decahedral or the lowest icosahedral minimum. We found that when a replica is pinned at the global minimum, it is easy for another replica to find the lowest icosahedral minimum. An example of this behavior is shown for LJ₇₅ in Fig. 2, which corresponds to a likely path as the system moves from A (the decahedral minimum) to B (the lowest icosahedral minimum). While for a LJ₃₈ cluster both paths (AB and BA) are sampled – although the rate of occurrence of AB is higher than that of BA by a factor of 5 – the path BA is no longer sampled for a LJ₇₅ cluster. This provides a clear example of ergodicity breaking due to the presence of two competing funnels in the PES. It appears that the difficulty in going from B to A is not so much associated to the energy barriers that must be crossed but to the fact that while many paths lead to B, paths leading from B to A are rare. In other words, nucleation of icosahedral order is much easier than nucleation of ccp order. There is no easy way to avoid this difficulty with optimization schemes that use the PES as the only objective

function. Indeed, all the schemes that we examined in this study, including our most efficient quantum sampling method, failed to find the global decahedral minimum of LJ₇₅ when starting from configurations that were in the catchment basin of the icosahedral minimum. Similar difficulties were encountered using alternative methods solely based on the PES [20].

A possible remedy is to bias the system with additional information other than that provided by the PES alone. For instance, one could add a penalty function that favors ccp order over icosahedral order. This could be expressed in terms of icosahedral and ccp order parameters. Alternatively, the same effect could be introduced by biasing the preselection of the MC moves. This is sometimes done in optimization schemes based on genetic algorithms [21]. By specifically constructing cluster structures that correspond to icosahedral and decahedral growth sequences, these schemes have been able to find good candidate global minimum structures for LJ clusters of up to 309 atoms [22]. Successful genetic algorithms for LJ clusters have also been reported by [23,24]. Another highly successful approach is the basin-hopping method in which SA is performed on the inherent PES rather than on the PES itself as we do here [17,25]. With this scheme all the minima for LJ clusters up to 110 atoms have been found [17]. Novel optimization strategies with a good success rate for large LJ clusters have been reported by [26,27].

In conclusion, the exponential growth of the number of local minima with cluster size appears to be an insurmountable difficulty for any optimization algorithm based solely on the knowledge of the PES. However, some algorithms may work on a more extended size

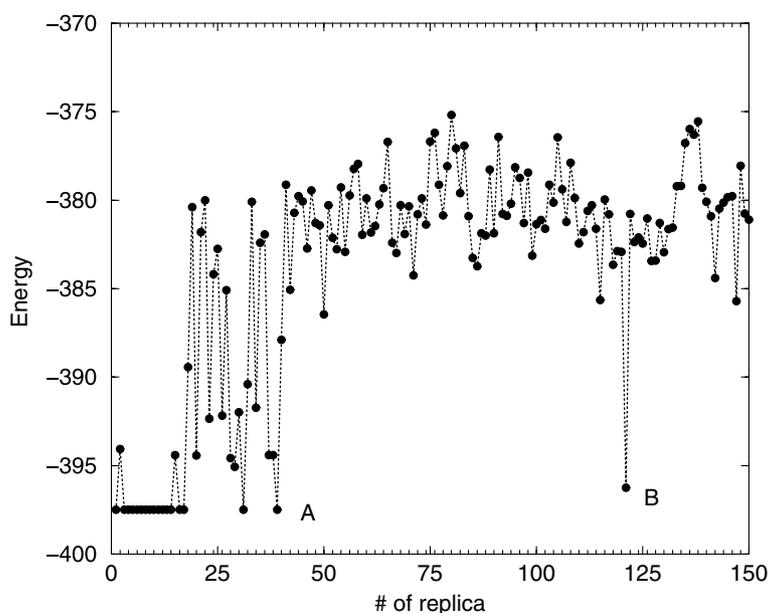


Fig. 2. Imaginary time path between global decahedral (A) and lowest icosahedral (B) minimum for LJ₇₅. This plot represents a snapshot of a run with $P = 500$, where the replica in the global minimum was held fixed.

range than others. We have shown here that quantum sampling methods are superior in this respect to their classical counterparts. In particular, we introduced a new approach, the RPQA method, which has been able to find the global minima for various clusters consisting of as many as 201 atoms. This approach may be useful in practical calculations on a variety of molecular systems as well as in more general optimization problems where the objective function depends on many variables and is characterized by the presence of many local minima.

Acknowledgement

This work was partially supported by the NSF through ITR grant CHE-0121432.

References

- [1] J.P.K. Doye, D.J. Wales, *Phys. Rev. Lett.* 80 (1998) 1357.
- [2] J.P.K. Doye, D.J. Wales, *J. Chem. Phys.* 109 (1998) 8143.
- [3] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, *Science* 220 (1983) 671.
- [4] J. Brooke, D. Bitko, T.F. Rosenbaum, G. Aeppli, *Science* 284 (1999) 779.
- [5] T. Kadowaki, H. Nishimori, *Phys. Rev. E* 58 (1998) 5355.
- [6] Y.-H. Lee, B.J. Berne, *J. Phys. Chem. A* 104 (2000) 86.
- [7] G.E. Santoro, R. Martonak, E. Tosatti, *R. Car*, *Science* 295 (2002) 2427.
- [8] R. Martonak, G.E. Santoro, E. Tosatti, *Phys. Rev. E* 057701 (2004) 70.
- [9] A.B. Finnila, M.A. Gomez, C. Sebenik, J.D. Doll, *J. Chem. Phys.* 80 (1984) 5709.
- [10] D.J. Wales, J.P.K. Doye, A. Dullweber, M.P. Hodges, F.Y. Naumkin, F. Calvo, J. Hernandez-Rojas, T.F. Middleton, The Cambridge Cluster Database. Available from: URL <<http://www-wales.ch.cam.ac.uk/CCD.html>>.
- [11] R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.
- [12] R.P. Feynman, *Statistical Mechanics*, Benjamin Cummings, Reading, 1972.
- [13] L.S. Schulman, *Techniques and Applications of Path Integration*, Wiley, New York, 1981.
- [14] D. Chandler, P.G. Wolynes, *J. Chem. Phys.* 74 (1981) 4078.
- [15] D.M. Ceperley, *Rev. Mod. Phys.* 67 (1995) 279.
- [16] M. Tuckerman, B.J. Berne, G.J. Martyna, M.L. Klein, *J. Chem. Phys.* 99 (1993) 2796.
- [17] D.J. Wales, J.P.K. Doye, *J. Phys. Chem. A* 101 (1997) 5111.
- [18] E. Marinari, G. Parisi, *Phys. Rev. Lett.* 75 (1995) 288.
- [19] F.H. Stillinger, T.A. Weber, *J. Chem. Phys.* 83 (1985) 4767.
- [20] M.D. Wolf, U. Landman, *J. Phys. Chem. A* 102 (1998) 6129.
- [21] D.M. Deaven, K.M. Ho, *Europhys. Lett.* 19 (1992) 451.
- [22] C. Barrón, S. Gómez, D. Romero, A. Saavedra, *Appl. Math. Lett.* 12 (1999) 85; D. Romero, C. Barrón, S. Gómez, *Comp. Phys. Commun.* 123 (1999) 87.
- [23] B. Hartke, *J. Comp. Chem.* 20 (1999) 1752.
- [24] Y. Xiang, H. Jiang, W. Cai, X. Shao, *J. Phys. Chem. A* 108 (2004) 3586.
- [25] D.J. Wales, H.A. Scheraga, *Science* 285 (1999) 1368.
- [26] X. Shao, H. Jiang, W. Cai, *J. Chem. Inf. Comput. Sci.* 44 (2004) 193.
- [27] L. Cheng, W. Cai, X. Shao, *Chem. Phys. Lett.* 404 (2005) 182.