



Note

Multi-scale global optimization of all-atom molecular models of n-alkanes

Rajeswar R. Gattupalli, Angelo Lucia*

Department of Chemical Engineering, University of Rhode Island, Kingston, RI 02881, United States

ARTICLE INFO

Article history:

Received 22 September 2008

Received in revised form 8 November 2008

Accepted 14 November 2008

Available online 25 November 2008

Keywords:

Multi-scale global optimization

Terrain/funneling method

Molecular conformation of solid n-alkanes

ABSTRACT

All-atom potential energy models of solid n-alkanes are important in understanding wax formation. Numerical results for molecular conformation computations using the multi-scale terrain/funneling method of Lucia and co-workers [Lucia, A., DiMaggio, P. A., & Depa, P. (2004a). A geometric terrain methodology for global optimization. *Journal of Global Optimization*, 29, 297; Lucia, A., DiMaggio, P. A., & Depa, P. (2004b). Funneling algorithms for multi-scale optimization on rugged terrains. *Industrial & Engineering Chemistry Research*, 43, 3770; Gattupalli, R. R., & Lucia, A. (2007). Molecular conformation of n-alkanes using terrain/funneling methods. *Journal of Global Optimization*, doi:10.1007/s10898-007-9206-5] are reported for all-atom potential energy models of the solid n-alkanes: dodecane, hexadecane, tetracosane, and triacontane, and compared with the basin hopping method of Wales and Doye [Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *Journal of Physical Chemistry A*, 101, 5111]. It is shown that the terrain/funneling method is capable of finding the global minimum energy molecular conformation for all examples in a reliable and efficient manner while basin hopping often fails. These numerical results provide strong evidence demonstrating that the multi-scale terrain/funneling method is a powerful global optimization method for solving molecular conformation problems.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction and motivation

In molecular modeling it is often important to be able to reliably find the lowest energy structure for a given molecule. However, this is generally difficult because of the very large number of minima and saddle points on the potential (or Gibbs) energy surface. Fortunately, this particular class of global optimization problems can be characterized by an objective function with two distinct geometries – a noisy or rugged landscape at small length scales and a funnel shaped or non-quadratic geometry at large length scales (see Onuchic, Luthey-Schulten, & Wolynes, 1997) – suggesting that multi-scale global optimization methods are ideally suited for molecular conformation problems.

The motivation for this work is our continuing interest in building molecular understanding of wax formation. Prior work by Gattupalli and Lucia (2009) applied the terrain/funneling method to liquid n-alkanes using united atom potential energy models. The results of these optimization calculations showed that the united atom model is unsuitable for solid n-alkanes and that global minimum energy conformations exhibit considerable wrapping at the ends of the molecule. In this note, we present numerical results for the multi-scale terrain/funneling method of global optimization of

Gattupalli and Lucia (2009) for all-atom potential energy models of solid n-alkanes. Accordingly this paper is organized in the following way. First we present a brief summary of the multi-scale terrain/funneling method and an overview of the basin hopping method of Wales and Doye (1997). Next numerical results for finding the global minimum for all-atom potential energy models of n-dodecane ($n\text{-C}_{12}\text{H}_{26}$), n-hexadecane ($n\text{-C}_{16}\text{H}_{34}$), n-tetracosane ($n\text{-C}_{24}\text{H}_{50}$), and n-triacontane ($n\text{-C}_{30}\text{H}_{62}$) are presented for the terrain/funneling method. We also compare our results with basin hopping and the experimentally observed structure of n-alkanes. This article ends with a summary of our numerical findings and a preview of future work.

2. A summary of the terrain/funneling method

The multi-scale global optimization method of Gattupalli and Lucia (2009) is an important refinement of the original terrain/funneling method of Lucia, DiMaggio, and Depa (2004a, 2004b), which consists of small-scale and large-scale optimization calculations to build both local and global representations of the objective function surface. At the small length scale, the terrain method of Lucia and co-workers (2002, 2003, 2004) and the mean value theorem are used to gather average gradient and Hessian matrix information while the funneling method (Gattupalli & Lucia, 2009; Lucia et al., 2004a, 2004b) is used at the large length scale to build funnel approximations and make large conformational moves

* Corresponding author. Tel.: +1 401 874 2814; fax: +1 401 874 4689.
E-mail address: lucia@egr.uri.edu (A. Lucia).

on the objective function surface. The key to the success of this multi-scale terrain/funneling method of global optimization lies in the communication of information between length scales. Average gradient and curvature information gathered by the terrain method at the small length scale is passed to the funneling method, which uses this average information to compute a new estimate of the funnel minimum. The funneling method then passes this new estimate of the funnel minimum to the terrain method in order to identify the region on the objective function landscape for the next set of small-scale explorations. The optimization calculations alternate between length scales until the value of the objective function can no longer be reduced during the funneling phase. As pointed out by Gattupalli and Lucia (2009), the novel and computationally important features of their recent modification of the original terrain/funneling method are that it

- (1) uses averaged and point-wise gradient and Hessian matrix information;
- (2) forces the objective function to decrease monotonically at the large length scale;
- (3) ensures that each funneling phase of the computations converges to a stationary point;
- (4) replaces point-wise gradients of zero at any computed local minimum or saddle point by non-zero average gradient information to avoid getting trapped at local minima. We refer the reader to Gattupalli and Lucia (2009) for the details of the terrain/funneling method and a detailed literature survey.

3. An overview of basin hopping

The basin hopping method of Wales and Doye (1997) is a stochastic/deterministic global optimization method because it uses a combination of a Monte Carlo method and a conjugate gradient method and touted to be a very reliable methodology. Basin hopping is often initialized using either a random or known molecular conformation which is quenched to a configuration with (local) minimum energy, E_{\min} , using the Polak-Ribiere variant of the conjugate gradient method with a loose convergence tolerance (e.g., $\|g\| < 0.1$, where g is the gradient of the potential energy function). This minimum configuration is then given a small Monte Carlo trial move to a new configuration, which is subsequently quenched. Sometimes convergence to the previously located minimum occurs while other times basin hopping converges to a new minimum energy conformation, E_{new} . Acceptance of any new minimum energy configuration is determined by the Metropolis scheme (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) using the Boltzmann factor $\exp[-\Delta E_{\min}/k_B T]$, where ΔE_{\min} is the energy difference between E_{new} and E_{\min} , k_B is the Boltzmann constant, and T is absolute temperature. This procedure of alternating between trial Monte Carlo moves and a quench step using the conjugate gradient method is repeated until some pre-specified number of Monte Carlo steps has been executed. The molecular conformation with the lowest energy is deemed the best (or global minimum) configuration. The details of basin hopping can be found in Wales and Doye (1997).

4. Experimentally observed conformation of n-alkanes

Results from X-ray diffraction experiments reported in the literature (see Norman & Mathisen, 1972; Nyburg & Gerson, 1992; Turner, 1971) clearly establish that the global minimum energy molecular conformation of solid n-alkanes is a planar, zigzag structure of carbon atoms with the hydrogen atoms on each methylene (CH_2) group located in planes passing through the carbon atoms perpendicular to the chain axes.

5. Numerical results

In this section, we present numerical results for determining the global minimum of all-atom potential energy models of the n-alkanes: n- $\text{C}_{12}\text{H}_{26}$, n- $\text{C}_{16}\text{H}_{34}$, n- $\text{C}_{24}\text{H}_{50}$ and n- $\text{C}_{30}\text{H}_{62}$ using the multi-scale terrain/funneling method and basin hopping. All numerical results were performed on a Dell High Precision 670 workstation using the LF95 Lahey-Fujitsu compiler. The convergence tolerance for the terrain/funneling method was $\|g\| < \varepsilon = 10^{-8}$, where $\|g\|$ denotes the two norm of a vector and g is the gradient of the potential energy. For basin hopping, 5000 Monte Carlo steps and a convergence tolerance of $\|g\| < 0.1$ in the quench (or conjugate gradient) phase were used. Our implementation of basin hopping was first used to compute global minima for several Lennard-Jones clusters. These computations showed that our basin hopping implementation matched results (i.e., approximately the same Monte Carlo step on which the global minimum was first computed by Wales and Doye) reported in Wales and Doye (1997). This, in our opinion, validates our implementation of the basin hopping method.

5.1. All-atom models of n-alkanes

The potential energy function used in this work consisted of bond length, bond angle, torsion angle and van der Waals energy contributions. Constraints to avoid rotational and translational singularities were also used. Details of the mathematical models can be found in Gattupalli (2008). Empirical force field parameters were taken from MacKerell et al. (1998).

Fig. 1 shows numerical results for multi-scale global optimization calculations using the terrain/funneling method for an all-atom representation of n- $\text{C}_{30}\text{H}_{62}$, which has 270 unknown Cartesian coordinates. Note that the terrain/funneling method easily finds the experimentally observed planar, zigzag molecular conformation of n- $\text{C}_{30}\text{H}_{62}$.

The details for the terrain/funneling computations are as follows. The calculations were initialized with two different pairs of starting points for triacontane, which all had very large potential energy values of $E = 10^7$ kcal/mol. Each pair of starting points was used in conjunction with the funneling algorithm to find a stationary point on the potential energy surface. The first stationary point

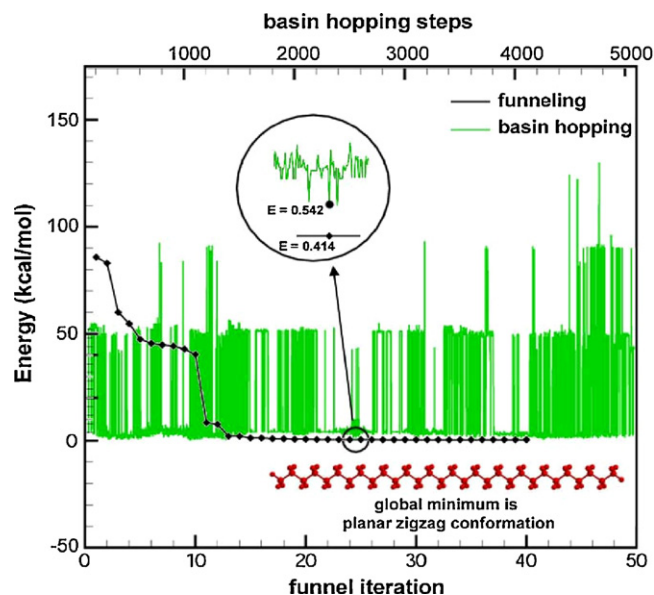


Fig. 1. Terrain/funneling calculations for triacontane using all-atom potential.

Table 1
Comparison of terrain/funneling and basin hopping methods for n-alkane conformation.

Problem	Terrain/funneling ^a		Basin-hopping ^b	
	Time ^c (h)	Function calls	Time ^c (h)	Function calls
Dodecane	4.97	26,737	23.22	985,000 ^d
Hexadecane	12.88	39,369	36.12	1,120,000 ^d
Tetracosane	18.52	18,638	160.45	1,662,000 ^d
Triacontane	64.00	40,132	190.24	1,922,361 ^d

^a Includes f, g and h evaluations.

^b BH: 5000 basin hopping steps; only f and g evaluations.

^c Time on a 3.6 GHz Dell 670 High Precision workstation.

^d Failed to find global minimum.

that was located was a local minimum and had a corresponding potential energy of $E_1 = 139.6909$ kcal/mol while the second stationary point was a saddle point and had a potential energy of $E_2 = 85.8561$ kcal/mol. Small-scale terrain calculations were then conducted to gather average gradient and average Hessian information along the terrain paths in the neighborhoods of these stationary points. The first set of terrain calculations required 21,684 function, gradient, and Hessian matrix evaluations and 118,805 s (33 h) of computer time while the second set of terrain calculations needed 18,363 function, gradient, and Hessian matrix evaluations and 111,601 s (31 h) of computer time.

Information from these small-scale optimization calculations was then used to initialize iterative large-scale funnel calculations and the funneling algorithm located a new stationary point corresponding to a global minimum on the energy surface at $E_3 = 0.414556$ kcal/mol in 40 funnel iterations and 35.76 s. Repeated small-scale calculations around E_3 and a subsequent set of funnel iterations gave the same minimum and thus the terrain/funneling method terminated with $E_4 = E_3 = 0.414556$ kcal/mol as the global minimum of the potential energy for triacontane.

Basin hopping was also used to determine the conformation of triacontane but failed to find the global minimum as shown by the inset in Fig. 1. The lowest energy structure found by the basin hopping method had a potential energy of $E = 0.542$ kcal/mol and required 1,922,361 function and gradient calls and 190.24 h of computer time for the 5000 Monte Carlo steps used. Many other starting points were tried and yielded the same general results. That is, the terrain/funneling method routinely found the global minimum molecular configuration while basin hopping failed.

5.2. A summary of numerical results

Table 1 gives a summary of the numerical results for the terrain/funneling and basin hopping methods of global optimization for all n-alkane molecular conformation examples studied in this work. Note that basin hopping failed on all example problems studied while the terrain/funneling algorithm successfully solved all problems. Moreover, all n-alkane structures calculated by the terrain/funneling algorithm exactly match experimentally observed structures of n-alkanes. As noted earlier, that molecular structure is a planar, zigzag structure of carbon atoms with the hydrogen atoms on each methylene (CH_2) group located in planes passing through the carbon atoms perpendicular to the chain axes.

6. Conclusions

The multi-scale terrain/funneling method of global optimization was used to determine the global minimum molecular conformation of all-atom potential energy models of solid n-alkanes ($n\text{-C}_{12}\text{H}_{26}$, $n\text{-C}_{16}\text{H}_{34}$, $n\text{-C}_{24}\text{H}_{50}$, and $n\text{-C}_{30}\text{H}_{62}$). In all cases, the calculated global minimum energy conformation was a planar zigzag structure and agrees with experimentally observed crystal structures. Our numerical results clearly demonstrate that the terrain/funneling methodology is a reliable, robust and efficient global optimization method and can be used for molecular conformation problems with as many as 300 degrees of freedom. The terrain/funneling method was also compared to the basin hopping method of Wales and Doye (1997). These comparisons showed that the terrain/funneling method is much faster and more reliable than basin hopping.

The next phase of this research will study the dynamics of phase transitions of n-alkanes, which are enabled by the multi-scale global optimization results for liquid n-alkanes presented in Gattupalli and Lucia (2009) and those given in this paper for solid n-alkanes. These results provide the groundwork for understanding of $R \rightarrow R$, $R \rightarrow \text{LO}$ and $\text{LO} \rightarrow \text{LO}$ transitions from a molecular perspective, where R is a rotator or liquid-like phase and LO denotes a low ordered temperature solid phase. These transitions play an important role in the waxing of petroleum fuels.

Acknowledgment

Acknowledgment is made to the Donors of the American Chemical Society Petroleum Research Fund for support of this research under Grant No. ACS PRF # 44793-AC9.

References

- Gattupalli, R. R. (2008). Advanced methods for global optimization. *Ph.D. Thesis*, University of Rhode Island.
- Gattupalli, R. R., & Lucia, A. (2009). Molecular conformation of n-alkanes using terrain/funneling methods. *Journal of Global Optimization*, 43, 429.
- Lucia, A., DiMaggio, P. A., & Depa, P. (2004a). A geometric terrain methodology for global optimization. *Journal of Global Optimization*, 29, 297.
- Lucia, A., DiMaggio, P. A., & Depa, P. (2004b). Funneling algorithms for multi-scale optimization on rugged terrains. *Industrial & Engineering Chemistry Research*, 43, 3770.
- Lucia, A., & Yang, F. (2002). Global terrain methods. *Computers & Chemical Engineering*, 26, 529.
- Lucia, A., & Yang, F. (2003). Multivariable terrain methods. *AIChE Journal*, 49, 2553.
- MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., et al. (1998). All-atom empirical potential for molecular modeling and dynamic studies of proteins. *Journal of Physical Chemistry B*, 102, 3586.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087.
- Norman, N., & Mathisen, H. (1972). The crystal structure of lower paraffins. *Acta Chemica Scandinavica*, 26, 3913.
- Nyburg, S. C., & Gerson, A. R. (1992). Crystallography of the even n-alkanes: Structure of $\text{C}_{20}\text{H}_{42}$. *Acta Crystallographica*, B48, 103.
- Onuchic, J. N., Luthey-Schulten, Z., & Wolynes, P. G. (1997). Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry*, 48, 545.
- Turner, W. R. (1971). Normal alkanes. *Industrial & Engineering Chemistry Product Research*, 10, 238.
- Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *Journal of Physical Chemistry A*, 101, 5111.