

This is the accepted manuscript made available via CHORUS. The article has been published as:

# Protein-folding simulations of the hydrophobic-hydrophilic model by combining pull moves with energy landscape paving

Jingfa Liu, Gang Li, and Jun Yu

Phys. Rev. E **84**, 031934 — Published 30 September 2011

DOI: [10.1103/PhysRevE.84.031934](https://doi.org/10.1103/PhysRevE.84.031934)

# Protein folding simulations of the hydrophobic-hydrophilic model by combining pull moves with energy landscape paving

Jingfa Liu<sup>1, 2\*</sup>, Gang Li<sup>3</sup> and Jun Yu<sup>4</sup>

1. School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing, 210044, China

2. Network Information Center, Nanjing University of Information Science & Technology, Nanjing, 210044, China

3. School of Mathematics & Physics, Nanjing University of Information Science & Technology, Nanjing, 210044, China

4. Department of Mathematics & Statistics, The University of Vermont, Burlington, Vermont, VT 05405, USA

The energy landscape paving (ELP) method is a class of heuristic global optimization algorithm based on Monte Carlo sampling. By incorporating generation of initial conformation based on greedy strategy, conformation update mechanism based on pull moves and some heuristic off-trap strategies into the improved ELP method, we propose a new version of ELP, called ELP-pull moves. We test ELP-pull moves on both two-dimensional (2D) and three-dimensional (3D) hydrophobic-hydrophilic (HP) protein folding model. For ten 2D benchmark sequences of length ranging from 20 to 100, the proposed algorithm finds the lowest energies so far. Within the achieved results, the algorithm converges more rapidly and efficiently than the previous methods. For all ten 3D sequences with length 64, the ELP-pull moves method finds new lower energies within comparable computational times. The numerical results demonstrate our algorithm is a powerful method to study lattice protein folding model.

## I. INTRODUCTION

Predicting a protein's tertiary structure from its primary amino acid sequence is one of most challenging problems in biology. Two major difficulties have been challenging researchers. One is the design of appropriate energy functions. The effective energy function can generally distinguish the native states from non-native states of protein molecule. The other is the exploration of the vast space of all possible structures. The latter has been suggested as the current bottleneck [1], and attracts many experts who work on computational problems in local and global optimization. In this paper, we address to the latter which can be investigated on a particular simplified model: the hydrophobic-hydrophilic (HP) lattice model [2] in both two and three dimensions, which is a widely studied abstract one and has been used by biochemists to evaluate new hypothesis of protein structure formation.

In order to search for the minimum energy conformations for HP model, people have presented many heuristic algorithms, including genetic algorithm (GA) [3], and its variations (improved genetic algorithm (IGA) with a new selection scheme and multiple-points crossover [4] and guided genetic algorithm (GGA) [5]), particle swarm optimization (PSO) [6], evolutionary Monte Carlo (EMC) [7], varieties of Monte Carlo (MC) [8], the hybrid of GA and tabu search (GTS) [9], elastic net algorithm with local search (ENLS) [10,11], gradient-directed Monte Carlo (GDMC) [12], GA based on optimal secondary structures (GAOSS) [13], and hybrid Taguchi genetic algorithm (HTGA) [14] that combines genetic algorithm, Taguchi method, and PSO. Meanwhile, some statistic approaches such as fragment regrowth via energy-guided sequential sampling

---

\* Corresponding author.  
E-mail address: jfliu@nuist.edu.cn



The HP lattice protein folding problem can be formally defined as follows: given an amino acid sequence  $s=s_1s_2\cdots s_n$ , we try to find an energy-minimizing conformation of  $s$ , that is, to find  $c^*\in T(s)$  so that  $E(c^*)=\min\{E(c) \mid c\in T(s)\}$ , where  $T(s)$  is the set of all the valid conformations of  $s$ . Even though the HP lattice model is highly simplified, it has been proven that the corresponding folding problem remains to be NP-complete [29].

### III. METHODS

#### A. Energy landscape paving

The energy landscape paving (ELP) method [22, 23] is an improved Monte Carlo (MC) global optimization method. As all good stochastic global optimizers, it is designed to explore low-energy conformations while avoiding at the same time entrapment in local minima. This is achieved by performing low-temperature MC simulations, but with a modified energy expression designed to steer the search away from regions that have already been explored. This means that if a conformation  $c$  is hit, the energy  $E(c)$  is increased by a “penalty”, and replaced by energy  $\tilde{E}(c) = E(c) + f(H(q, t))$ , where the “penalty” term  $f(H(q, t))$  is a function of the histogram  $H(q, t)$ . In the present paper, we choose  $\tilde{E}(c) = E(c) + kH(E, t)$  as the replacement of the energy  $E$ , where  $H(E, t)$  is the histogram in energy at MC sweep  $t$  and  $k$  is a constant. During the process of execution of ELP, we update the histogram by adding 1 to the frequency of the corresponding bin if the energy  $E(c)$  of the newly generated conformation  $c$  falls into a certain bin, where a “bin” denotes an entry of the histogram, and all bins in the histogram are the same in size. We set the size of every bin  $E_{bin}=1$  in HP model. The sampling weight for a conformation  $c$  is defined as  $\omega(\tilde{E}(c)) = \exp(-\tilde{E}(c)/k_B T)$ , where  $k_B T$  is the thermal energy at the (low) temperature  $T$ , and  $k_B$  is Boltzmann constant.

ELP has been applied successfully to rough energy landscapes for finding low-energy conformations in the off-lattice protein folding problems [23-25] and the circular packing problems [26,27]. In ELP minimization, the sampling weight of a local minimum conformation decreases with the time that system stays in that minimum, and consequently the probability to escape the minimum increases. The accumulated histogram function  $H(E, t)$  from all previously visited energy at the MC sweeps helps the simulation escape local entrapments and surpass high-energy barrier easier. However, in ELP, there exists a technical flaw [26,27]. From the current conformation  $c_1$ , we consider an attempted move in ELP which yields a new lower energy minimum  $c_2$  that has never been visited before and energy of which happens to fall into the same bin which contains other energies previously visited in the earlier steps. Undesirably, the likelihood of accepting this new energy minimum  $c_2$  becomes small as the result of the penalty term  $kH(E, t)$ , i.e., the ELP search may miss this new lower energy conformation  $c_2$  near  $c_1$ . To overcome this shortcoming, Refs.[26] and [27] give an improvement on ELP. In the improved ELP, the acceptability of the new conformation  $c_2$  is determined by a comparison between  $E(c_1)$  and  $E(c_2)$ , where two cases are possible: (a)  $E(c_2) < E(c_1)$  and (b)  $E(c_2) \geq E(c_1)$ . For

case (a), the simulation unconditionally accepts the new conformation  $c_2$  and starts a new round of iteration; for case (b), if the new conformation  $c_2$  satisfies the condition expression  $r(0,1) < \exp\{[\tilde{E}(c_1, t) - \tilde{E}(c_2, t)]/k_B T\}$ , where  $r(0,1)$  denotes a random number between 0 and 1, then the simulation accepts  $c_2$  and starts a new round of iteration, otherwise does not accept  $c_2$  and restores  $c_1$  as the current conformation.

### B. Generation of initial conformation based on greedy strategy

For the HP protein folding problem in two and three dimensions, we build 2D and 3D Cartesian coordinate system, respectively. The first amino acid and the second amino acid are put at (0, 0) and (0, 1), respectively, for two dimensions, and at (0, 0, 0) and (0, 0, 1) for three dimensions. Subsequently the  $i$ th ( $3 \leq i \leq n$ ) amino acid is first “pseudo-placed” at every position which is adjacent to the  $(i-1)$ th amino acid and not occupied by other amino acids, where “pseudo-place” means the  $i$ th amino acid is placed temporarily and will be removed from the corresponding position after computing the energy of the partial conformation which consists of the previous  $i-1$  amino acids placed and the  $i$ th amino acid. Then we put the  $i$ th amino acid at the position, where the energy of the corresponding partial conformation is lowest. This process repeats until a conformation with  $n$  amino acids is produced.

### C. Conformation update based on pull moves

In ELP, each MC step must update current conformation. We use a local move set (i.e. pull moves of Refs.[30] and [31]) as the conformation update move set. The set of pull moves is complete, reversible, and local [30], which makes them efficient for the conformation update. Any valid conformation is connected to all others by pull moves. Therefore, pull moves determine how other amino acids respond in HP lattice model when one amino acid is moved in one direction.

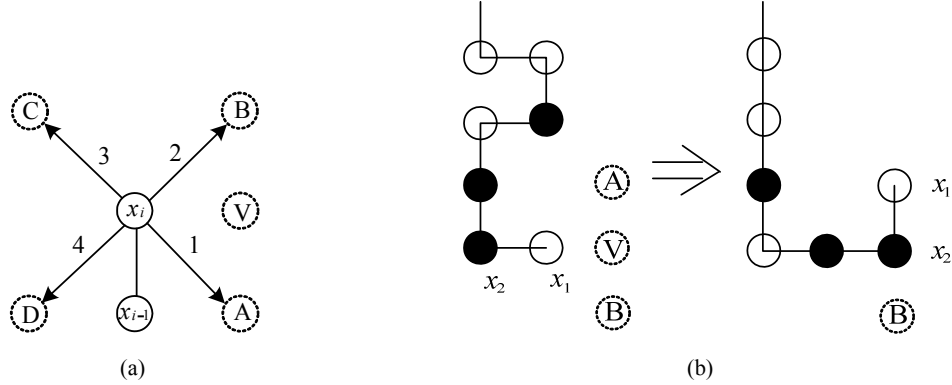


FIG. 2. Pull moves as the conformation update. (a) Four move directions of amino acid  $i$  ( $2 \leq i \leq n-1$ ); (b) End moves of amino acid  $i$  ( $i=1$ ).

First, we briefly describe the set of pull moves on 2D square lattice, i.e. a single  $XY$  plane. According to the pull move rules, four move directions, i.e. 1, 2, 3, and 4 are defined in Fig. 2. If the position in one direction (such as A, B, C, D) is already occupied by the other amino acid, the move in that direction is not valid. If

position A, B, C, or D is vacant, the moves to these positions may be permitted. Suppose the grid points  $x_{i-1}$  and  $x_i$  to be occupied by the  $(i-1)$ th and  $i$ th amino acids. Consider the pull moves of the  $i$ th amino acid. If the grid point A is unoccupied, label the fourth grid point in the mini-square including  $x_{i-1}$ ,  $x_i$  and A as V (see Fig. 2(a)). If  $V=x_{i+1}$  (which is occupied by the  $(i+1)$ th amino acid), then the pull move simply operates by moving  $x_i$  to A and generates a new conformation. If both V and A are unoccupied, the pull moves then operates by moving  $x_i$  to A,  $x_{i+1}$  to V,  $x_{i+2}$  to  $x_i$ ,  $x_{i+3}$  to  $x_{i+1}$ , and so on, until a new legal conformation is reached by the least number of grid point moves. In the pull moves described above the amino acids are pulled one by one in descending order. By symmetry the amino acid positions can also be pulled in ascending order in a pull move. The first and last amino acids in the chain require special pull moves [30]. In Fig.2 (b), if position A or B is vacant, the moves to these positions may be permitted. Let A and V be two adjacent unoccupied grid points. V is adjacent to the first amino acid position  $x_1$ . The end move displaces  $x_1$  to A,  $x_2$  to V,  $x_3$  to  $x_1$ ,  $x_j$  to  $x_{j-2}$ , and so on, until a new legal conformation is reached by the least number of grid point moves (Fig. 2(b)). By symmetry, the end move on the last amino acid is similarly defined.

In ELP, to update the current conformation, we consider in turn pull moves of every amino acid in the chain. Starting from the first amino acid, we execute pull moves for all legal move positions of each amino acid. The  $i$ th ( $1 \leq i \leq n$ ) amino acid may at most be moved to the four diagonal positions (see Fig.2). First, we “pseudo-move” it to every legal position. Then we complete remaining moves by pull move rules until a new legal conformation is reached. After computing the energy of the corresponding conformation for each legal position, we move the  $i$ th amino acid to the position, where the energy of the conformation obtained by pull moves is lowest. This process is repeated until a new conformation is accepted (the acceptance criterion is shown in section A) or pull moves on all  $n$  amino acid are executed.

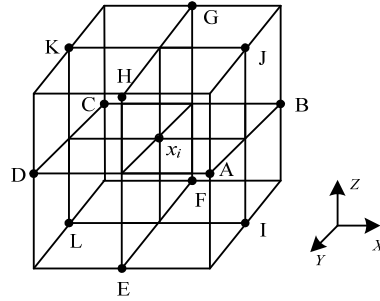


FIG. 3. Twelve potential diagonal move positions (i.e. indices A, B, C, D, E, F, G, H, I, J, K, and L) of the amino acid  $i$  ( $1 \leq i \leq n$ ) in 3D lattice.

For 3D cubic lattice, the pull moves of a grid point can be similarly defined as above. It is noteworthy that in 3D lattice an initial step of one pull move within one of the three planes may induce subsequent moves within the other planes. In addition, in 2D lattice, the  $i$ th ( $1 \leq i \leq n$ ) amino acid (which occupies grid point  $x_i$ ) may at most be moved to the four diagonal positions (see Fig.2 (a)), but in 3D lattice, at most to the twelve diagonal positions (see Fig. 3).

#### D. The off-trap strategy

When we apply pull moves to update conformation in ELP, it is possible to each new conformation obtained by pull moves is not accepted. This means that the search is trapped into local minima that may locate at narrow and deep valleys of the energy landscape. For jumping out of local minima, we execute the following three moves in turn: 1)  $90^\circ$  rigid rotations (Fig.4 (a)); 2) parallel moves (Fig.4 (b)); 3) crankshaft moves (Fig.4 (c)). In  $90^\circ$  rigid rotations, we suppose that grid point  $x_{i-1}$  (which is occupied by amino acid  $i-1$ ) is a pivot and the grid points  $x_i, x_{i+1}, x_{i+2}, \dots, x_n$  compose a rigid body. The whole rigid body rotates  $90^\circ$  according to grid point  $x_{i-1}$  along the  $XY$  plane for 2D lattice (see Fig. 4 (a)), and **along** one of three planes  $XY, XZ$ , and  $YZ$  for 3D lattice, gaining an updating conformation. In parallel moves, the grid points  $x_i, x_{i+1}, x_{i+2}, \dots, x_n$  compose a rigid body, and the whole rigid body **parallelly** moves along  $x$ -axis or  $y$ -axis direction for 2D lattice, and  $x$ -axis,  $y$ -axis or  $z$ -axis for 3D lattice until grid point  $x_i$  is still adjacent to  $x_{i-1}$  (see Fig.4 (b)). In crankshaft moves, a  $u$ -shaped structure consisting of four connected neighbors in the sequence rotates  $180^\circ$  for 2D lattice, and  $90^\circ$ ,  $180^\circ$ , or  $270^\circ$  for 3D lattice according to the axis consisting of the grid points  $x_{i-1}$  and  $x_{i+2}$  (see Fig. 4(c)). If these three moves can not produce a valid conformation, we restore the previous conformation  $\bar{c}$  of the current conformation  $c$  as the new current conformation and continue a new round of iteration of ELP.

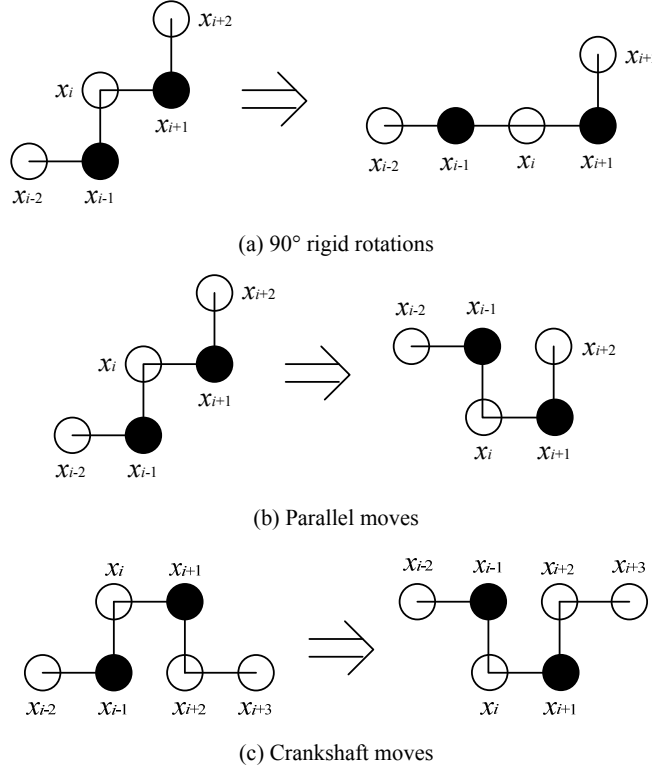


FIG. 4. Three off-trap strategies in HP model

#### E. Description of algorithm

By incorporating the generation of initial conformation based on greedy strategy, conformation update mechanism based on pull moves and some heuristic off-trap strategies into the improved ELP method, a new







GAOSS, MC-pull moves, GDMC-pull moves and ELP-pull moves find conformations with energy of -42. Three quite different conformations found by ELP-pull moves are shown in Fig. 5. In Fig. 5(a) and (c) there exist  $\alpha$ -helix,  $\beta$ -sheet,  $\beta$ -turn, and their mixture secondary structures, while in Fig. 5(b), there are mainly  $\alpha$ -helix secondary structures. For sequence 2D9(85-mer), two out of seven methods, GDMC-pull moves and ELP-pull moves find the ground state conformations with the lowest free energy, -53, but SISPER, GAOSS and MC-pull moves find the energy of -52, and other two methods, MC without pull moves and GA do not report the results. One of the ground state conformations by ELP-pull moves for sequence 2D9 is shown in Fig. 6. For sequence 2D10(100-mer), three out of seven methods, SISPER, GDMC-pull moves, and ELP-pull moves find conformations with energy of -48, and MC without pull moves, GA and GAOSS do not report the results of this instance, whereas MC-pull moves finds the energy of -47. One of the ground state conformations by ELP-pull moves for sequence 2D10 is shown in Fig. 7. From Figs. 5-7, one can see that all of the conformations possess a compact hydrophobic core.

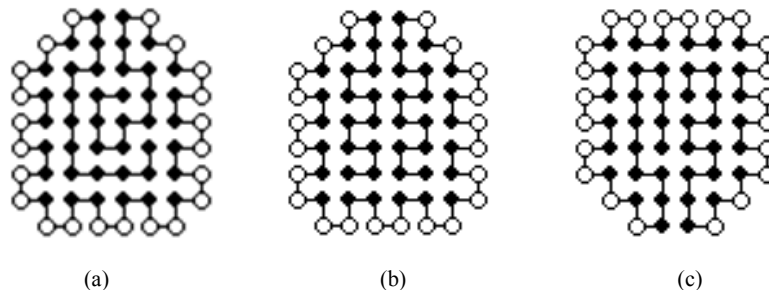


FIG. 5. Three conformations with energy of -42 for sequence 2D8 (the length-64 sequence) found by ELP-pull moves.

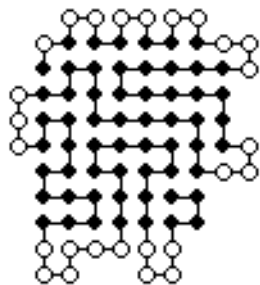


FIG. 6. The conformation with energy of -53 for sequence 2D9 (the length-85 sequence) found by ELP-pull moves.

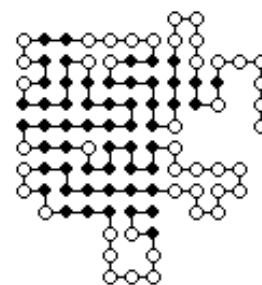


FIG. 7. The conformation with energy of -48 for sequence 2D10 (the length-100 sequence) found by ELP-pull moves.

In addition, the number of valid conformations scanned in ELP-pull moves for each sequence is also listed in Table III, in comparison with the results by MC [3], GA [3], MC-pull moves [12], and GDMC-pull moves [12]. Table III shows that ELP-pull moves spends less time to obtain the lowest energy than these four methods. There is no information about time cost for SISPER [18] and GAOSS [13], so we cannot make such a comparison. From Table III, one can also see that the putative ground state energies obtained by GDMC-pull moves and ELP-pull moves for all ten 2D sequences are the same, but the ELP-pull moves method requires much less time to obtain them than GDMC-pull moves. Therefore, ELP-pull moves explores the conformation surfaces more efficiently than MC, GA, SISPER, GAOSS, MC-pull moves and GDMC-pull moves.

TABLE IV. Comparison of performances of different methods on ten 3D HP sequences listed in Table II. NA means data not available. The number in each cell is the minimum energy obtained by the corresponding method for the respective HP sequence. The numbers in parentheses are the numbers of valid conformations scanned before the lowest energy values are found.

3D Seq.	MC <sup>a</sup>	GA <sup>a</sup>	IGA <sup>b</sup>	GGA <sup>c</sup>	PSO <sup>d</sup>	ELP-pull moves <sup>e</sup>
3D1	-12(4 139 486)	-27(2 119 775)	-22	NA	-28(1 131 552)	-31(102 640)
3D2	-17(4 086 574)	-29(2 286 289)	-26	-30(14 186)	-31(456 877)	-36(209 922)
3D3	-24(3 958 530)	-35(1 831 102)	-35	-38(493)	-39(113 315)	-44(101 422)
3D4	-18(4 077 468)	-34(2 315 112)	NA	-36(27 450)	-36(1 730 129)	-39(1 007 703)
3D5	-20(4 027 596)	-32(2 040 915)	NA	NA	-38(1 602 646)	-41(737 062)
3D6	-16(4 114 480)	-29(2 160 690)	NA	-30(1 899)	-31(410 586)	-34(405 219)
3D7	-15(4 128 584)	-20(2 317 862)	NA	NA	-27(1 296 319)	-28(51 015)
3D8	-19(4 067 513)	-29(2 391 876)	NA	NA	-35(1 113 330)	-37(87 188)
3D9	-19(4 053 207)	-32(2 121 287)	NA	-34(14 848)	-35(404 199)	-39(255 703)
3D10	-14(4 125 584)	-24(2 287 394)	NA	-25(362)	-27(175 053)	-31(111 953)

<sup>a</sup>Values are from Ref. 3.

<sup>b</sup>Values are from Ref.4.

<sup>c</sup>Values are from Ref.5.

<sup>d</sup>Values are from Ref.6.

<sup>e</sup>Values are from this current work.

For ten 3D sequences with 64-mers [3], we list the results of the ELP-pull moves algorithm in Table IV, in comparison with those by Monte Carlo (MC) [3], genetic algorithm (GA) [3], improved genetic algorithm [IGA] with a new selection scheme and multiple-points crossover [4], guided genetic algorithm (GGA) [5], and particle swarm optimization (PSO) [6]. From Table IV one can see that ELP-pull moves finds new lower free energies than these five methods for all ten 3D sequences. Figs. 8-17 show typical conformations of these putative lowest energy states. It is obvious that each of these conformations possesses a compact hydrophobic core. Moreover, ELP-pull moves scans less the number of valid conformations to obtain the lowest free energy than MC [3], GA [3], and PSO [6] for every sequence. In addition, one can notice that GGA [5] spent less times to obtain their results for sequences 3D2, 3D3, 3D4, 3D6, 3D9, and 3D10 than ELP-pull moves, but the lowest energies by GGA are far high (at least a difference of -3) from those by our method. For other four sequences, we cannot make a comparison between ELP-pull moves with GGA since GGA does not report the corresponding results.

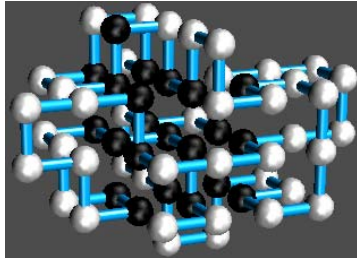


FIG. 8. (Color online) Typical conformation with  $E=-31$  of sequence 3D1.

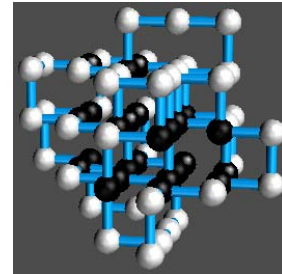


FIG. 9. (Color online) Typical conformation with  $E=-36$  of sequence 3D2.

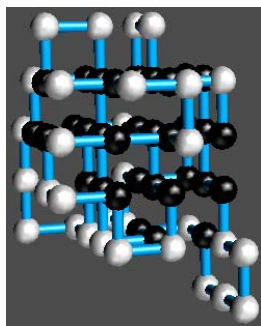


FIG. 10. (Color online) Typical conformation with  $E=-44$  of sequence 3D3.

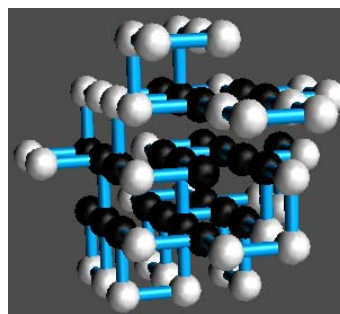


FIG. 11. (Color online) Typical conformation with  $E=-39$  of sequence 3D4.

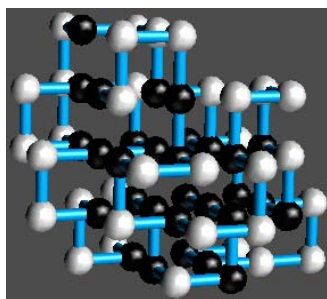


FIG. 12. (Color online) Typical conformation with  $E=-41$  of sequence 3D5.

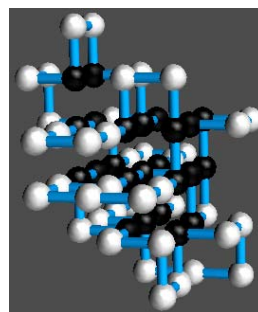


FIG. 13. (Color online) Typical conformation with  $E=-34$  of sequence 3D6.

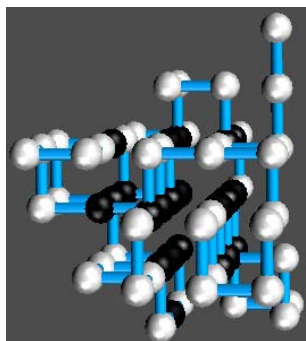


FIG. 14. (Color online) Typical conformation with  $E=-28$  of sequence 3D7.

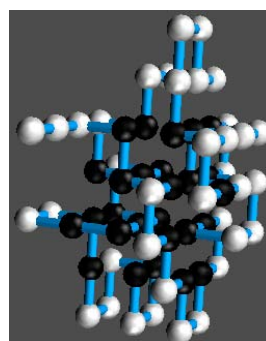


FIG. 15. (Color online) Typical conformation with  $E=-37$  of sequence 3D8.

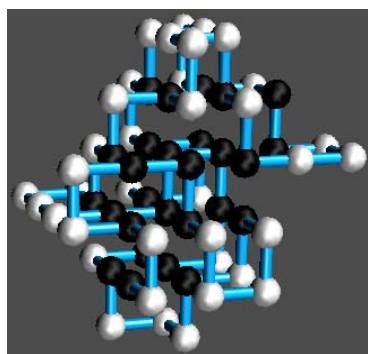


FIG. 16. (Color online) Typical conformation with  $E=-39$  of sequence 3D9.

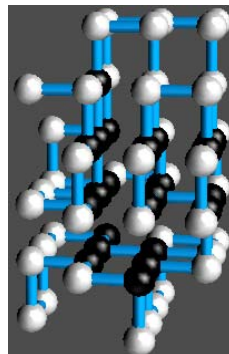


FIG. 17. (Color online) Typical conformation with  $E=-31$  of sequence 3D10.

## V. CONCLUSIONS

Stochastic global optimization method ELP has been applied successfully to rough energy landscape of continuous space for finding low-energy conformations, including the off-lattice protein folding problems, the circular packing problems, etc. However, few researches applied ELP to discrete combinatorial optimization problem. In this paper, to demonstrate the efficiency of ELP in discrete space, the protein folding problem in both 2D and 3D HP lattice model is studied, and the performance of the ELP procedure for each case is described. In ELP, we introduce pull moves to execute local search and update conformation. Because the landscape for protein folding is rugged, and the set of pull moves is local, ELP may trap the optimization in some local minima that locate at narrow and deep valleys of the energy landscape. So, some heuristic off-trap strategies are used to jump out of the local minima.

Our results indicate that for the 2D model the ELP method combined pull moves (ELP-pull moves) explores the conformation surface more efficiently than MC, GA, SISPER, GAOSS, MC-pull moves and GDMC-pull moves, and find the known lowest energies obtained by previous researchers in all cases. Moreover, for all ten 3D sequences, the ELP-pull moves algorithm outperforms MC, GA, IGA, GGA, and PSO, and finds new lower energies within comparable computational times. Briefly, this paper presents a new more efficient algorithm for finding lower free energy states of 2D and 3D HP lattice proteins. It is not hard to see that the proposed method is easy to be extended to other discrete optimization problem. Furthermore, our study suggests that a proper combination of normal stochastic global optimization method and local search can reduce the cost of stochastic global search and enhance the efficiency of search of the algorithm. It could be an efficient mechanism to construct high-performance algorithm in a certain problem.

## ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Jiangsu Province (BK2010570), the China Postdoctoral Science Foundation funded Project (20100471350), Jiangsu Planned Projects for Postdoctoral Research Funds (1001030B), Natural Science Foundation of Education Committee of Jiangsu Province (09KJB520008), the National Public Benefit Research Foundation of China (GYHY200906006), and Qing Lan Project.

## REFERENCES

- [1] P. Bradley, K. M. Misura and D. Baker, *Science* **309**, 1868 (2005).
- [2] K. F. Lau and K. A. Dill, *Macromolecules* **22**, 3968 (1989).
- [3] R. Unger and J. Moult, *J. Mol. Biol.* **231**, 75 (1993); R. Unger and J. Moult, in: *Proceedings of the Fifth International Conference on Genetic Algorithm*, San Francisco, CA, 581 (1993).
- [4] F. L. Custódio, H. J. C. Barbosa and L. E. Dardenne, *Genet. Mol. Biol.* **27**, 611 (2004).
- [5] M. T. Hoque, M. Chetty and L. S. Dooley, in: *Proceedings of 2006 IEEE Congress on Evolutionary*

- Computation, Vancouver, BC, Canada, 2339 (2006).
- [6] F. Kanj, N. Mansour, H. Khachfe and F. Abukhzam, in: Proceedings of ACS/IEEE International Conference on Computer Systems and Application, Rabat, Morocco, 732 (2009).
- [7] F. Liang and W. H. Wong, J. Chem. Phys. **115**, 3374 (2001).
- [8] U. Bastolla, H. Frauenkron, E. Gerstener, P. Grassberger and W. Nadler, Proteins: Struct., Funct., Genet. **32**, 52 (1998).
- [9] T. Jiang, Q. H. Cui and G. H. Shi, J. Chem. Phys. **119**, 4592 (2003).
- [10] Y. Z. Guo, E. M. Feng and Y. Wang, J. Chem. Phys. **125**, 154102 (2006).
- [11] Y. Z. Guo and E. M. Feng, J. Chem. Phys. **125**, 234703 (2006).
- [12] X. Q. Hu, D. N. Beratan and W. T. Yang, J. Chem. Phys. **131**, 154117 (2009).
- [13] C. H. Huang, X. B. Yang and Z. H. He, Comput. Biol. Chem. **34**, 137 (2010).
- [14] C. J. Lin and M. H. Hsieh, Expert Syst. Appl. **36**, 12446 (2009).
- [15] J. F. Zhang, S. C. Kou and J. S. Lin, J. Chem. Phys. **126**, 225101 (2007).
- [16] S. C. Kou and J. Oh, J. Chem. Phys. **124**, 244903 (2006).
- [17] H. Y. Lu and G. K. Yang, Comput. Math. Appl. **57**, 1855 (2009).
- [18] J. L. Zhang and J. S. Liu, J. Chem. Phys. **117**, 3492 (2002).
- [19] P. Grassberger, Phys. Rev. E. **56**, 3682 (1997).
- [20] H. P. Hsu, V. Mehra, W. Nadler and P. Grassberger, J. Chem. Phys. **118**, 444 (2003); H. P. Hsu, V. Mehra, W. Nadler and P. Grassberger, Phys. Rev. E. **68**, 021113 (2003).
- [21] W. Q. Huang and Z. P. Lü, Chin. Sci. Bull. **49**, 2092 (2004); W. Q. Huang and Z. P. Lü, Phys. Rev. E. **72**, 016704 (2005).
- [22] U. H. E. Hansmann and L. T. Wille, Phys. Rev. Lett. **88**, 068105 (2002).
- [23] J. F. Liu and W. Q. Huang, J. Theor. Comput. Chem. **5**, 587 (2006).
- [24] A. Schug, W. Wenzel and U. H. E. Hansmann, J. Chem. Phys. **122**, 194711 (2005).
- [25] J. F. Liu, S. J. Xue, D. B. Chen, H. T. Geng and Z. X. Liu, J. Biol. Phys. **35**, 245 (2009).
- [26] J. F. Liu and G. Li, Sci. China Inf. Sci. **53**, 885 (2010).
- [27] J. F. Liu, S. J. Xue, Z. X. Liu and D. H. Xu, Comput. Indust. Eng. **57**, 1144 (2009).
- [28] K. A. Dill, Biochemistry **24**, 1501 (1985).
- [29] B. Berger and T. Leight, J. Comput. Biol. **5**, 27 (1998).
- [30] N. Lesh, M. Mitzenmacher and S. Whitesides, in: Proceedings of the 7<sup>th</sup> Annual International Conference on Research in Computational Molecular Biology, Berlin, Germany, 188 (2003).
- [31] A. A. Albrecht, A. Skaliotis and K. Steinhöfel, Comput. Biol. Chem. **32**, 248 (2008).