J Mol Model (2013) 19:3883–3891
DOI 10.1007/s00894-013-1907-8

ORIGINAL PAPER

Enhanced hybrid search algorithm for protein structure prediction using the 3D-HP lattice model

Changjun Zhou · Caixia Hou · Qiang Zhang · Xiaopeng Wei

Received: 20 November 2012 / Accepted: 30 May 2013 / Published online: 4 July 2013

Abstract The problem of protein structure prediction in the hydrophobic-polar (HP) lattice model is the prediction of protein tertiary structure. This problem is usually referred to as the protein folding problem. This paper presents a method for the application of an enhanced hybrid search algorithm to the problem of protein folding prediction, using the three dimensional (3D) HP lattice model. The enhanced hybrid search algorithm is a combination of the particle swarm optimizer (PSO) and tabu search (TS) algorithms. Since the PSO algorithm entraps local minimum in later evolution extremely easily, we combined PSO with the TS algorithm, which has properties of global optimization. Since the technologies of crossover and mutation are applied many times to PSO and TS algorithms, so enhanced hybrid search algorithm is called the MCMPSO-TS (multiple crossover and mutation PSO-TS) algorithm. Experimental results show that the MCMPSO-TS algorithm can find the best solutions so far for the listed benchmarks, which will help comparison with any future paper approach. Moreover, real protein sequences and Fibonacci sequences are verified in the 3D HP lattice model for the first time. Compared with the previous evolutionary algorithms, the new hybrid search algorithm is novel, and can be used effectively to predict 3D protein folding structure. With continuous development and changes in amino acids sequences, the new algorithm will also make a contribution to the study of new protein sequences.

Keywords 3D HP lattice model · Crossover and mutation · Particle swarm optimizer · Tabu search · Protein folding prediction

C. Zhou (✉) · C. Hou · Q. Zhang · X. Wei
Key Laboratory of Advanced Design and Intelligent Computing, Dalian University, Dalian, China
e-mail: zhou-chang231@163.com

Introduction

Proteins—polymers composed of chains of amino acids—are present in many different types, and play many roles, in living organisms. Initially, a protein is a linear chain of amino acids. Proteins fold, under the influence of several chemical and physical factors, into the three-dimensional (3D) structures that determine their biological functions and properties [1–3]. In most successful protein structure prediction (PSP), essentially a lattice model has been utilized for folding backbone sampling at the top of a hierarchical approach [4]. Given the amino acid sequence of a protein, the prediction of that protein's tertiary structure [5] is known as the protein folding problem. The prediction of protein structure is a NP-hard (nondeterministic polynomial time) problem, which attracts many researchers to study this area.

The prediction of protein structure is one of the most prominent problems in computational biology [6–8]. The function of a protein depends mainly on its tertiary structure, which in turn depends on its primary structure. It is known that ill-formed proteins (due to wrong folding) are the origin of numerous diseases, such as cystic fibrosis, Parkinson's disease, Alzheimer's disease, and some types of cancer [9]. If we can predict the tertiary structures of proteins with high accuracy, it will become easy to treat these diseases. Due to their great importance for medicine and biochemistry, proteins have been the focus of much research [10], with the result that much information is now available. Research on 3D structure prediction is very important in the search for new drugs with specific functionality.

Despite the growing number of proteins already discovered, further research is still ongoing. The growing calculation power of computers can be used to solve the complex computational problems in protein structure prediction. In this regard, there are two major challenges to solve [11]:

(1) finding good measures to verify the qualities of candidate structures;
(2) given such measures, determining optimal or close-to-optimal structures for a given amino acid sequence.

A variety of mathematical models for protein have been put forward, of which the AB off-lattice model (Fig. 1), which also uses only two types of monomers, called "A" (hydrophobic) and "B" (hydrophilic), and hydrophobic-polar (HP) lattice model (Fig. 2), are the most popular.

In recent years, many algorithms for protein folding research have been proposed in the 3D HP lattice model, e.g., the PERM, MOSE, SISPERS, Evolutionary [13], and Genetic algorithms, etc. In this paper, we present an enhanced hybrid search algorithm that combines an enhanced particle swarm algorithm with an enhanced tabu search (TS) algorithm. The enhanced particle swarm optimizer (PSO) algorithm [14] appends the operation of crossover (single-point and two-point crossover), and the enhanced TS algorithm adds the operation of mutation.

The remainder of the paper is organized as follows. Section "HP lattice model" gives an introduction to the HP lattice model; "Calculation of free energy" denotes the function of energy and describes computing approaches; "Relevant algorithm introduction" summarizes relevant algorithms used in the hybrid algorithm; "MCMPSO-TS hybrid search algorithm" describes the MCMPSO-TS algorithm in detail. Experimental results obtained by our method and by other methods are compared in the section on "Experimental results and discussion". A final "Conclusions" section concludes the paper.

HP lattice model

The HP model is based on the observation of the hydrophobic interaction between two amino acid residues. This is the driving force for protein folding and the development of a native protein state [15–18]. The simplest model for the study of protein folding is known as the HP model, which exists in either 2D or 3D versions [19]. The 3D lattice model proposed by Dill [5, 20, 21] is applied in this paper.

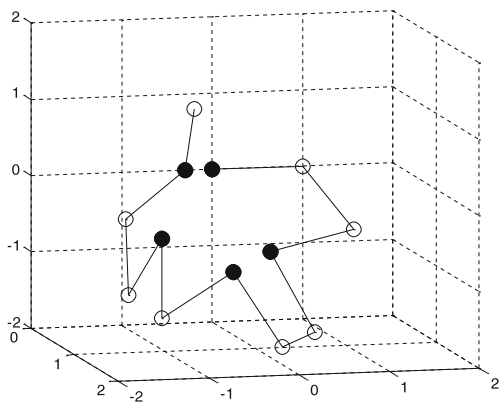

Fig. 1 AB off-lattice model [12]

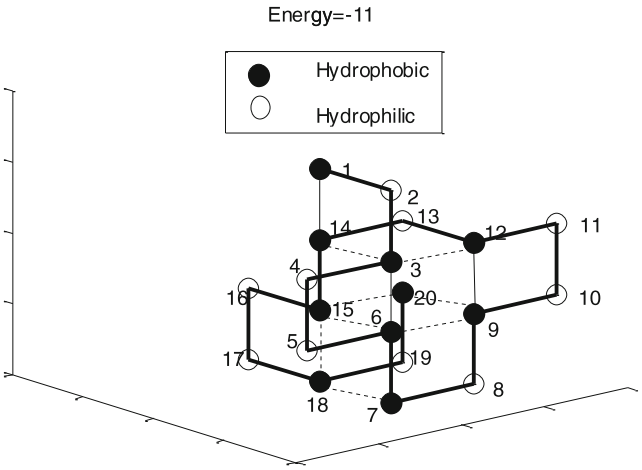

Fig. 2 3D HP lattice model [11]

The simplified HP lattice model divides the amino acids into two different types: hydrophobic (H) and hydrophilic (P). The proposed lattice model has the following two advantages [11]:

(1) amino acids are modeled as single beads rather than by every atom;
(2) the beads are restricted to a rigid lattice.

Based on the HP lattice model, the energy of a given conformation is defined as the number of connections between Hs of topological neighbors that are not direct neighbors in the sequence [19]. Generally speaking, natural proteins adopt both the lowest energy structure and the most stable. The energy function determines the interactions, and different types of amino acids interact differently.

Calculation of free energy

The free energy of amino acids can be calculated by the following formula [20]:

$$\varepsilon_{ij} = \begin{cases} -1.0 & \textit{the pair of } H \textit{ and } H \textit{ residues} \\ 0.0 & \textit{others} \end{cases} \quad (1)$$

The free energy of the protein sequence can be obtained as follows:

$$E = \sum_{i,j} \triangle r_{ij} \varepsilon_{ij} \quad (2)$$

where the parameter

$$\triangle r_{ij} = \begin{cases} 1 & s_i \text{ and } s_j \text{ are adjacent but not connected amino acids} \\ 0 & \text{others} \end{cases} \quad (3)$$

The HP lattice model has two bead types: H (hydrophobic) and P (hydrophilic). In the sequences of amino acids, H–H connections are assigned an energy value of −1. The free energy is minimum while the number of H–H connections is maximum. The optimal conformation in the HP model is the one that has the maximum number of H–H connections, which gives the lowest energy value.

The problem of optimization of protein folding can be transformed into the calculation of the minimum free energy of protein folding conformation. Formally, given an HP sequence $s=s_1,s_2,s_3,\cdots,s_n$, find a construction of s with minimum energy. That is: find $c^* \in C(s)$, such that $E(c^*)=min\{E(c)|c \in C\}$, where $C(s)$ is the set of all valid conformations for s [6, 11, 22, 23].

An introduction to the relevant algorithms

Enhanced particle swarm optimizer

The PSO algorithm was first put forward by Eberhart and Kennedy [24]. The algorithm is based on the information transmission in bird foraging behaviors; each individual can remember not only the best position that it has found currently, but also the best position that individuals find globally [24]. Using these two optimal values, the birds can find food.

In the PSO, the trajectory of each individual in the search space is adjusted by dynamically altering the velocity of each particle, according to its own experience and the experience of neighbor particles in the search space. In the D-dimensional search space, every particle is a point. If there are m particles, the position of ith particle is denoted as $X_i=(x_{i1},x_{i2},\cdots,x_{iD})$, the current velocity is defined as $V_i=(v_{i1},v_{i2},\cdots,v_{iD})$.

According to the fitness function, the best position of each particle is as follows:

$$P_i = (p_{i1},p_{i2},p_{i3},\cdots,p_{iD})$$

The fittest particle found so far is

$$P_g = \left(p_{g1},p_{g2},p_{g3},\cdots,p_{gD}\right)$$

[11].

The velocity and position of the ith particle are updated by the formulae:

$$v_{(t,t+1)} = wv(i,t) + c_1 r_1 \left(pbest_{(i,t)} - x_{(i,t)}\right) + c_2 r_2 \left(gbest_{(g,t)} - x_{(i,t)}\right) \quad (4)$$

$$x_{(i,t+1)} = x(i,t) + v_{(i,t+1)}) \quad (5)$$

Here w is referred to as the inertia. Where c_1 and c_2 are referred to self confidence and swarm confidence respectively. r_1 and r_2 are random real numbers between 0 and 1.

Compared with standard PSO, the enhanced PSO has some prominent differences. Though updated formulas of positions and velocities of particles are unchanged, the parameter is updated by a new formula as follows:

$$w = w\_max - time^*(w\_max - w\_min)/MaxDT \quad (6)$$

The w_max and w_min are the range of inertia w, time is the circular times, *MaxDT* is the maximum number of iterations that make sure it is easy to find the suitable value and promote the efficiency. The greatest difference is that a multiple crossover operation, including single-point crossover and two-point crossover is used. In the enhanced PSO algorithm, the crossover operation is applied many times during the selection of the best solution and the search procedure.

Enhanced tabu search optimizer

Tabu search (TS)—a global neighboring search algorithm proposed by Glover—is the simulation of human memory function [25, 26]. The PSO algorithm converges slowly in the latter period, and sinks easily into local optimal solution [27]. TS is the best choice to solve this problem. The greatest difference between standard TS and enhanced TS is that mutation is applied to the standard TS search algorithm several times, which promotes the efficiency of the search and changes the search position.

MCMPSO-TS hybrid search algorithm

The new algorithm is composed of enhanced PSO and enhanced TS. The new algorithm is called MCMPSO-TS (multiple crossover and mutation PSO-TS).

Although PSO has many advantages, the disadvantages of poor convergence precision and the tendency to trap into the local optimum cannot be ignored. TS, as an extension of local neighboring search, is a global search algorithm. So, in this paper, we introduce TS to cover the shortcomings of PSO [11]. At the same time, we introduce the technologies of

crossover and mutation, which are very useful for improving the quality of the search procedure.

In this paper, we use the enhanced PSO output result as the input for the enhanced TS algorithm. In the search process, we utilize the operations of crossover and mutation to guarantee the search space. Further, we introduce a new encoded mode. The following sections explain the new hybrid search algorithm in detail.

Initialization

If the input of the amino acid sequence is composed of H and P, which is of length *n*, then each individual in the population is a string of length *n*. For example, HHPPHHPHPH is a string of length 20. If the symbol is H, the encoding is 1, otherwise 0, and the amino acid sequence can be written as 1100110101. The position of the amino acid is represented by a 3D coordinate. Since folding direction between two amino acids is the basic element in the protein 3D space structure, the folding direction is also an essential element to show the protein 3D space structure in the computer.

For the 3D lattice model, there are six folding directions, which are denoted by six figures in this paper as follows [28]:

‘1’ denotes positive direction of x-axis
‘2’ denotes negative direction of x-axis
‘3’ denotes positive direction of y-axis
‘4’ denotes negative direction of y-axis
‘5’ denotes positive direction of z-axis
‘6’ denotes negative direction of z-axis

In other words, in this paper, the amino acid is presented by a 4D coordinate. The first position is initialized as (0,0,0,0), which means the coordinate of the amino acid is (0,0,0). The first amino has no folding direction, the value is 0. The next folding direction of amino acid is random. For example, the sequence is HHPPHHPHPH, it is composed of 10 amino acids. So the sequence can be written as 1100110101. The position of the first amino acid is (0,0,0,0); it has no folding direction. The direction of the second amino acid can be produced at random. If it is 3 (the positive direction of the y-axis), the coordinate of the amino acid can be decided by the previous coordinate (0,0,0,0).

The previous coordinate (0,0,0,0) means:

amino(1,0)=0 coordinate of x-axis
amino(1,1)=0 coordinate of y-axis
amino(1,2)=0 coordinate of z-axis
amino(1,2)=0 folding direction

So the second is:

amino (2,0)=amino (1,0)=0 coordinate of x-axis
amino (2,1)=amino (1,1)+1=1 coordinate of y-axis
amino (2,2)=amino (1,2)=0 coordinate of z-axis
amino (2,3)=3 folding direction

In conclusion, the coordinate of second amino acid is (0,1,0,3).

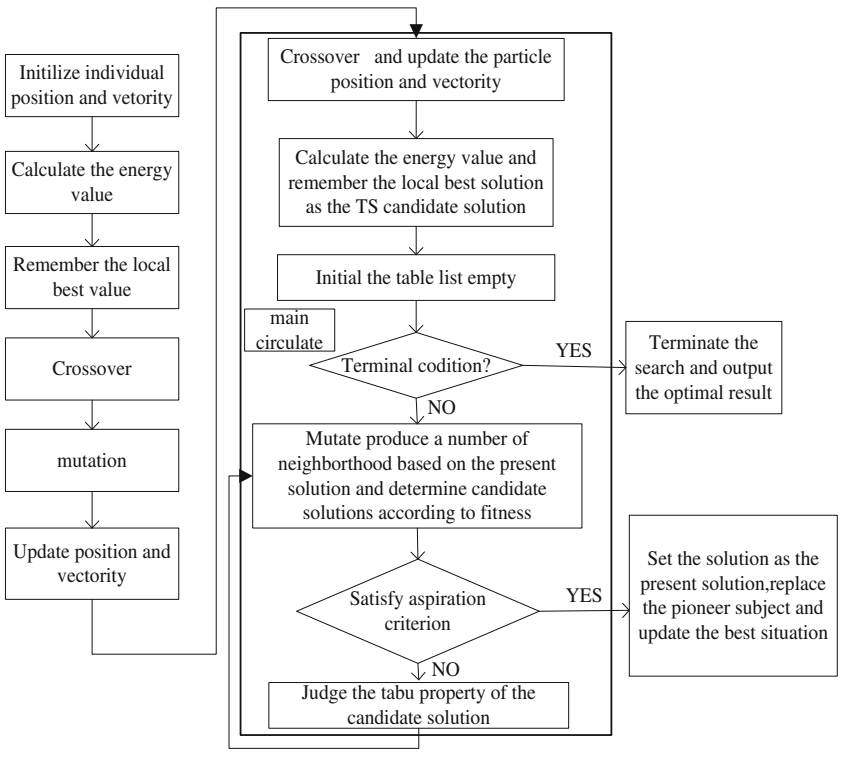

Fig. 3 Flowchart of multiple crossover and mutation particle swarm optimizer and tabu search (MCMPSO-TS) algorithm

Methods of crossover and mutation

In the paper, single-point crossover and two-point crossover are used. The method that is applied to the paper is the following: first, choose the individual that needs to be adopted for the operation of crossover; then produce the positions of the crossover randomly; at last swap the directions that are in the positions chosen. A single-point crossover is used to swap the directions from the position that is chosen to last. The two-point crossover is to swap the directions between the positions that are chosen. When the directions are changed, we should change the coordinates, too. Therefore, the method can generate population diversity.

Method of generating the neighborhood search

In this paper, we use the idea of mutation to generate a neighborhood solution. The disturbance mutation method is used to generate neighbor solutions of the current solutions. For the present solution, we utilize multiple-point (two-point in this paper) mutation in the early stage of the search to guarantee better diversity and single-point mutation in the later stage to assure convergence of the algorithm. The implementation is presented as follows:

We select the position of mutation x^k at random and use the following function to update it [29]:

$$x_{new}^{k} = floor\left(x^{k} + 2 \times \pi \times f(r) \times c \times rate^{i}\right) \tag{7}$$

Where c is a random number between 0 and 1, *rate* is the factor of scale and i, which changes from 0 to *NL-1* (*NL* is the size of the neighborhood), denotes the iterations generating the neighborhood solution. In this paper, we set *rate*=0.95 as in [30]. Coefficient $f(r)$ is defined as follows:

$$f(r) = \begin{cases} 0, & r \geq 0.5 \\ -1, & r < 0.5 \end{cases} \tag{8}$$

Here, r is a random number between 0 and 1.

Implementation procedure

This section describes the implementation of the MCMPSO-TS hybrid algorithm as follows:

Step 1: Initialize the parameter

Give the parameter c_1 and c_2, the range of w, the scale of particle, the iterations, and the length of tabu.

Step 2: Initialize the velocity and position

Using the method of generation of coordinate that is introduced above, initialize the position and velocity. The position cannot be the same.

Step 3: Calculate the fitness value

(1) calculate the fitness value by using the Eqs. (2) that is introduced in calculation of free energy
(2) sort the fitness values from small to large;
(3) choose the smallest one as the candidate.

Step 4: Crossover operation, update the positions and velocities of the particles

Step 5: Main algorithm

For time=1:*MaxDT*

(1) repeat the crossover procedure;
(2) update the positions and velocities of the particles using Eqs. (4) and (5);
(3) calculate the fitness value;
(4) initialize the tabu list and obtain the candidate solution;

while *runcount*<*T*

runcount=*runcount*+1;

(5) mutate and use Eq. (7) to update the position of mutation;
(6) calculate the fitness value;
(7) sort the value of energy from the smallest to the largest and choose the smallest one as *candidate_next*;

if *candidate_next*<*best_so_far.value*

(8) update *candidate_now* with the solution of neighborhood whose the value of energy is the smallest;
(9) put *candidate_now.xulie* into the *tabulist* then update the *best_so_far* with *candidate_now*;

else

for *n*=1:10

(10) *index*=*find*(*cat*(1,*neighorhood.value*)==*candidate_next*(*n*));

tx=*neighorhood*(*index*(1)).*xulie*; *tc*=*neighorhood*(*index*(1)).*value*;

(11) estimate that whether this object is in *tabulist* or not. If it is, update the *candidate_now* then put it into *tabulist*;

return *best_so_far*;

Step 6: Find the best solution and output it

Figure 3 shows the flowchart of the new algorithm:

Table 1 Running results of Fibonacci sequences. *E* Energy value, *MCMPSO-TS* multiple crossover and mutation particle swarm optimizer and tabu search algorithm

Length	Sequence	E(MCMPSO-TS)
5	HPPHP	−1
8	PHPHPPHP	−2
13	HPPHPPHPHPPHP	−5
21	PHPHPPHPHPPHPPHPHPPHP	−8
34	HPPHPPHPHPPHPPHPHPPHP HPPHPPHPHPPHP	−19

Experimental results and discussion

In this paper, the hybrid algorithm was implemented by MATLAB R2009b under a Windows XP system. The parameters were set as follows: MaxDT = 1,000. The number of particles was from 100 to 300. Run-count was 500. c_1 and c_2 are both 2. All the results are achieved by running many times and getting the best solution. The paper proposes the computation of Fibonacci sequences of amino acids by using MCMPSO-TS hybrid search algorithm in the 3D lattice model. At the same time, sequences of real proteins were also computed by the new hybrid search algorithm in the HP lattice model. Additionally, we also compared the results with those of published techniques.

Results of Fibonacci sequences

A Fibonacci sequence, which is defined as follows, was used in the protein folding prediction [24]:

$$S_0 = H, S_1 = P, ..., S_{i+1} = S_{i-1} * S_i \quad (9)$$

Where '*' is connection symbol. So S_2=*HP*, S_3=*PHP*, let H denotes hydrophobic while P denotes hydrophilic amino acids [24].

Table 1 shows the lowest energy values of Fibonacci sequences gained by the algorithm.

From Table 1, we can see that the results reflect the lowest energy values of Fibonacci sequences that were acquired by the new hybrid algorithm. The Fibonacci sequence was firstly clearly proposed in the 3D lattice model. The Fibonacci sequence is a classical mathematical model. It can reflect the energy of protein amino acids that adhere to some regular rules. Based on the point, we propose the sequence and put it into practice, which proves the value of research into the difference between the general sequences of protein amino acids and the sequences of protein amino acids that have some regular rules in energy and stability. This will be our next research goal. Figures 4 and 5 in the Appendix give the conformation of several sequences.

Results for real sequences

Next, we describe a real protein sequence. The sequence was downloaded from the PDB database (http://pdbbeta.rcsb.org/pdb/Welcome) [31]:

Table 2 Running results of real protein sequence

Name	Length	E(MCMPSO-TS)
1BXL	16	−6
1EDP	17	−4
1AGT	38	−19

- 1BXL GQVGRQLAIIGDDINR
- 1EDP CSCSSLMDKECVYFCHL
- 1AGT GVPINVSCTGSPQCIKPCK DQGMRFGKCMNRKCHCTPK

In this sequence the I,V,L,P,C,M,A,G are hydrophobic, and D,E,F,H,K,N,Q,R,S,T, W,Y are hydrophilic.

Table 2 lists the results from real protein sequences computed by the new algorithm. Because current technology cannot simulate more than a few micro-seconds of protein behavior, it is not possible to determine complete atomic detail [20, 33, 35] and the real protein cannot be simulated in a computer operation; however, the energy of the sequences of protein amino acids can be calculated based on the HP model. Although the results cannot reflect the structure of a real protein completely and we cannot see the approximate simulating structure of the real protein, the energy value gives us some reference for estimating the range of energy values of real protein structure and stability. This represents great progress. Figures 6 and 7 in the Appendix give the conformation of 1BXL and 1EDP.

Compared result

To obtain an obvious comparative result, this paper selects the same sequence as those examined in the 3D HP lattice model [11, 14, 28, 35]. The sequences and results are presented in Tables 4, 5 and 6 in the Appendix.

From Table 4, we can see that the results from the new algorithm are the same as those obtained currently, thus showing that the MCMPSO-TS algorithm is feasible. We can obtain a stable structure using the new algorithm. The new algorithm lays a solid basis for further research in this field, and represents a new method in this area. The ever increasing number of sequences DNA sequences available meant that the number of known amino acid sequences is also changing and increasing. The algorithm presented here may play an important role in promoting the development of protein structure prediction and provide a new train of thought for the research in this area, and may prove a valuable addition to the field. Figures 8 and 9 give conformation of S1 and S5 (see Appendix).

Tables 5 and 6 present sequences and energy values from [14]. From Table 6, we can see that our algorithm can achieve the best value of most sequences. However, for sequences longer than 48, the new algorithm is no longer superior. Nevertheless, there is no doubt that one more algorithm has been added to the field of protein structure prediction and that our new method enriches the database of algorithms. The proposed algorithm gives good reference values for further research, and is thus worth pursuing and continuing. Our results show that the new algorithm is better for computing the short amino acid sequences, but not long sequences. This will be the area we should continue to study next.

Conclusions

In this paper, we have modeled the protein structure prediction problem as a multimodal optimization problem. To foster its development, several multimodal optimization techniques have been implemented and tested. To summarize the findings of this work: first, we propose a new method and prove it to be feasible; second, real protein sequences and Fibonacci sequences were applied to the 3D lattice model; third, this new algorithm can achieve optimal values and conformations quickly for most short sequences.

In particular, the paper proposes the MCMPSO-TS algorithm for the problem of prediction of protein structure using the 3D HP lattice model. The new hybrid search algorithm includes enhanced PSO and enhanced TS, which use crossover and mutation separately, and finds a better energy value for most of the given sequences. Our algorithm is feasible. The proposed experimental method using Fibonacci sequences and real protein sequences is novel and represents a great step forward. Overall, the results are good and promise to support the continuity of the work. We believe that this paper makes a useful contribution to this area of research. We will focus on larger proteins, long sequences and the results of lower energy in future work.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No.31170797,30870573,61103057), the Program for Changjiang Scholars and Innovative Research Team in University (No.IRT1109), the Key Project of Chinese Ministry of Education (No.211036), the Program for Liaoning Excellent Talents in University (No.LR201003).

Appendix

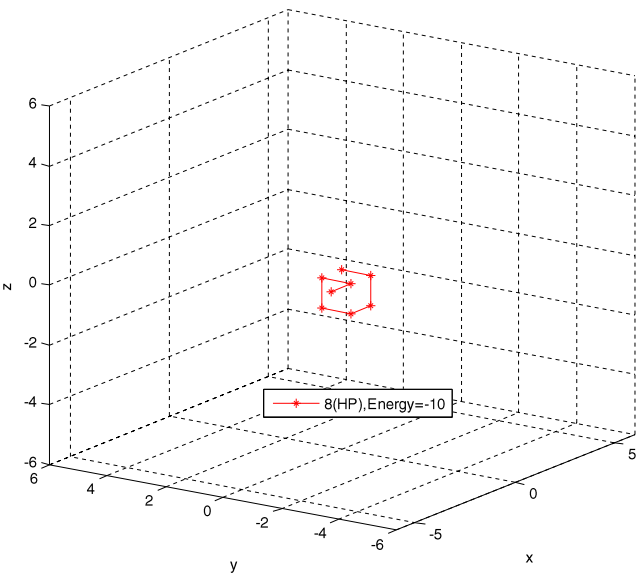

Fig. 4 Conformation of sequence of length 8

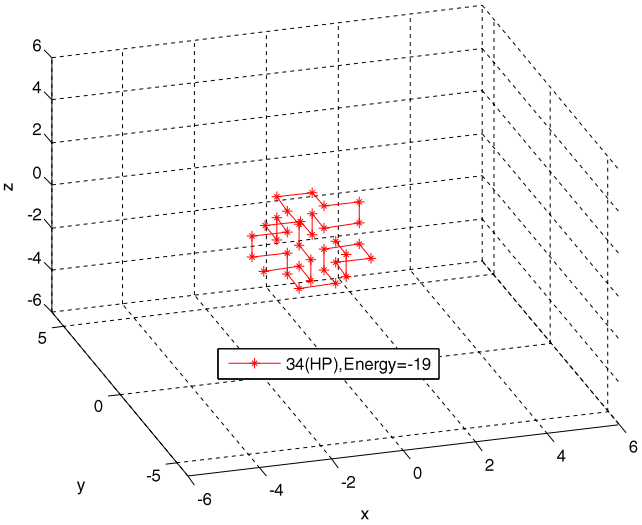

Fig. 5 Conformation of sequence of length 34

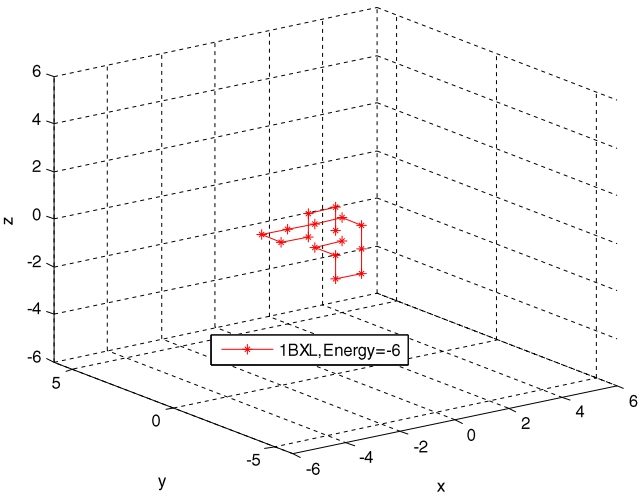

Fig. 6 Conformation of sequence of 1BXL

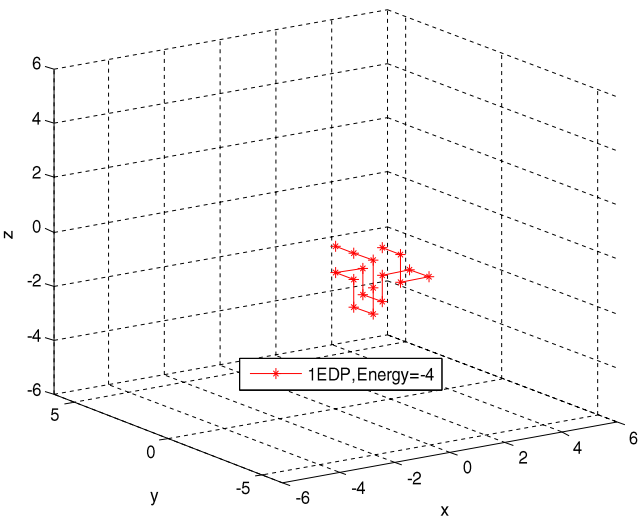

Fig. 7 Conformation of sequence of 1EDP

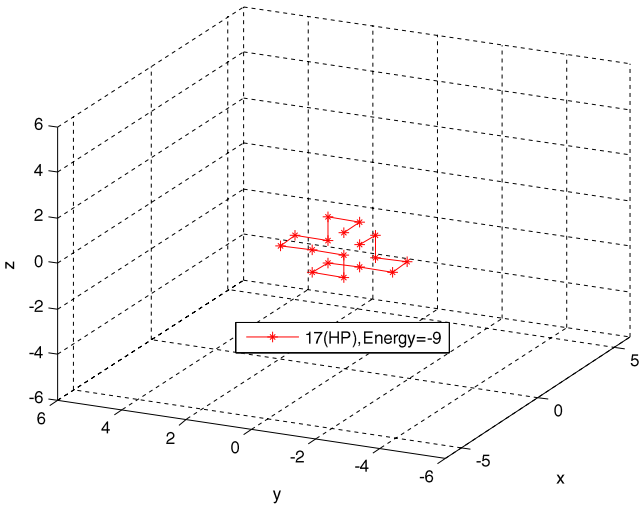

Fig. 8 Conformation of sequence of S1

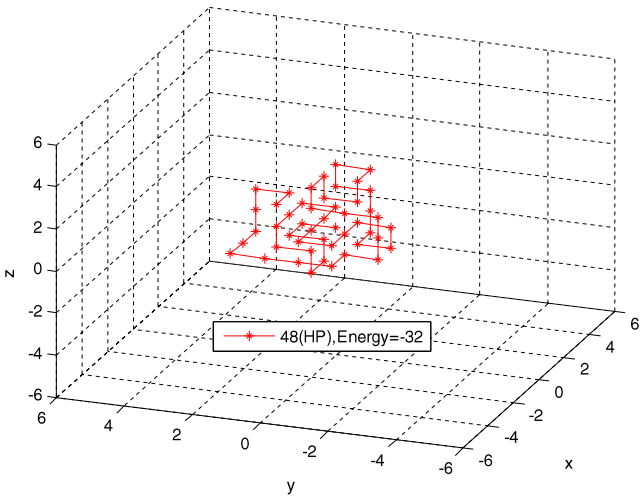

Fig. 9 Conformation of sequence of S5

Table 3 Sequences of comparison

Name	Length	Sequence
S1	17	HHHHPPHHHHHHHPPPH [28]
S2	20	HPHPPHHPHPPHPHHPPHPH [11]
S3	25	PPHPPHHPPPPHHPPPPHHPPPPHH [11]
S4	36	PPPHHPPHHPPPPPHHHHHHHPPHHPPPPHHPPHP [11]
S5	48	HPHHPPHHHHPHHHPPHHPPHPHHHPHPHHPPHHPPPHPPPPPPPPHH [35]

Table 4 Results of comparison

Name	E(HGA-SA) [28]	E(HGA-PSO) [11]	E(GA) [35]	E(MCMPSO-TS)
S1	−9	–	–	−9
S2	–	−11	–	−11
S3	–	−9	–	−9
S4	–	−18	–	−18
S5	–	–	−32	−32

Table 5 Sequences of length 27

Name	Sequence
1	PHPHPHHHPPHPHPPPPPPPPPPHHP
2	PHHPPPPPPPPPPHHPPHHPPHPPHPH
3	HHHHPPPPPHPPPPPHHHPPPPPPPH
4	HHHPPHHHHPPPHPHPPHHPPHPPPHH
5	HHHHPPPPHPHHPPPHHPPPPPPPPP
6	HPPPPPPHPHHHPPHHPPPHPPPPHPH
7	HPPHPHHPPPHPPPPPHPHHPHPHPHH
8	HPPPPPPPPPPPHPHPPPPPPPPHPHH
9	PPPPPPPHHHPPPHPHHPPPHPPHPPP
10	PPPPPHHPHPHPHPHPPHHPHHPHPPP

Table 6 Results of comparison of length 27

Name	E (Jouhnson et al.)	E (Version1)	E (Version4)	E (Version7)	E(MCMPSO-TS)
1	−9	−9	−7	−6	−8
2	−10	−10	−7	−7	−6
3	−8	−8	−7	−5	−8
4	−12	−12	−8	−7	−12
5	−8	−8	−7	−5	−8
6	−11	−12	−8	−7	−10
7	−13	−13	−10	−8	−11
8	−4	−4	−4	−3	−4
9	−7	−7	−6	−5	−7
10	−11	−11	−9	−7	−11

References

1. Mansour N, Kanj F, Khachfe H (2011) Enhanced genetic algorithm for protein structure prediction based on the HP model. Search Algorithms Appl 3:69
2. Soriano-Ursua MA, Correa-Basurto J, Trujillo-Ferrara JG et al (2011) Homology model and docking studies on porcine adrenoceptor: description of two binding sites. J Mol Model 17:2525–2538
3. Bishop OT, Kroon M (2011) Study of protein complexes via homology modeling, applied to cysteine proteases and the protein inhibitors. J Mol Model 17:3163–3172

4. Yang X, Huang C, He Z (2010) Protein folding simulations of 2D HP model by the genetic algorithm based on optimal secondary structures. Comput Biol Chem 34(137)
5. Rother K, Rother M, Boniecki M et al (2011) RNA and protein 3D structure modeling: similarities and differences. J Mol Model 17:2325–2336
6. Srivastava M, Gupta SK, Abhilash PC et al (2011) Structure prediction and binding sites analysis of curcin protein of Jatropha curcas using computational approachies. J Mol Model 18:2971–2979
7. Chakravarty S, Ghersi D, Sanchez R (2011) Systematic asssessment of assuracy of comparative model of protein belonging to different structural fold classes. J Mol Model 17:2831–2837
8. Curco D, Michaux C, Roussel G et al (2012) Stochastic simulation of structural properties of natively unfolded and denatured proteins. J Mol Model 18:4503–4516
9. Benitez CMV, Lopes HS (2009) A parallel genetic algorithm for protein folding prediction using 3D-HP side chain model. IEEE Congress on Evolutionary Computation, 1297, Trondheim, 18–21 May 2009
10. Subashini M, Devarajan PV (2011) Molecular dynamics simulation of drug uptake by polymer. J Mol Model 17:1141–1147
11. Lin C-J, Shih-Chieh S (2011) Protein 3D HP model folding simulation using a hybrid of genetic algorithm and particle swarm optimization. Int J Fuzzy Syst 13:140–141
12. Zhang X, Cheng W (2008) Protein 3D structure prediction by improved Tabu Search in Off-Lattice AB model. IEEE 187
13. Guo H, Lv Q, Wu J et al (2009) Solving 2D HP protein folding problem by parallel ant colonies. IEEE 2
14. Mansour N, Kanj F, Khachfe H (2012) Particle swarm optimization approach for protein structure prediction in the 3D HP model. Intediscip Sci Comput Life Sci (4):190–280
15. Wang M, Wang J (2011) A computerized protein-protein interaction modeling study of ampicillin antibody specificity in relation to biosensor development. J Mol Model 17:2873–2882
16. Durham E, Dorr B, Woetzel N et al (2009) Solvent accessible surface area approximations for rapid and accurate protein structure prediciton. J Mol Model 15:1093–1108
17. Li T, Wang J, Li Y et al (2011) Structure of the complex between Mucor pusillus pepsin and the key domain of *k*-casein for site-directed mutagenesis: a combined molecular modeling and docking approach. J Mol Model 17:1661–1668
18. Mazzucco TN, Zanconato S, De Lucrezia D et al (2011) Design and dynamic simulation of minimal metallo-proteins. J Mol Model 17:2919–2925
19. Zhou H, Lv Q (2009) A study on applying particle swarm optimization algorithm. Su Zhou College 15
20. Chira C, Horvath D (2011) Evolutionary algorithm for protein structure prediction in lattice models. Analele Unibersitatii de Vest, vol 1, p 9
21. Saraswathi S, Fernandez-Martinez JL, Kolinski A et al (2012) Fast learning optimized prediction methodology (FLOPRED) for protein secondary structure prediction. J Mol Model 18
22. Yang Y, Liu H, Juan D et al (2011) A combined molecular modeling study on a series of pyrazole/isoxazole based human Hsp90 *a* inhibitors. J Mol Model 17:3241–3250
23. Su S-C, Lin C-J, Ting C-K (2010) An effective hybrid of hill climbing and genetic algorithm for 2D triangular protein structure prediction. Proteome Sci 18:3
24. Chen X, Lv M, Zhao L et al (2011) An improved particle swarm optimization for protein folding prediction. IEEE 1:3–5
25. Glover F (1989) Tabu search. J Comput 1:190–206
26. Glover F (1990) Tabu search. J Comput 2:4–32
27. Zhang X, Wang T, Luo H et al (2010) 3D Protein structure prediction with genetic tabu search algorithm. BMC Syst Biol 4:3–4
28. Wang X, Chen D (2006) The research of protein configuration forecasting based on mixed genetic algorithm. Wuhan University of Technology, pp 39–40
29. Bui TN, Sundarra G (2005) An efficient genetic algorithm for predicting protein tertiary structures in the 2D HP model. Computer Science Program 385
30. Zhang XL, Cheng W (2008) An improved tabu search algorithm for 3D protein folding problem, PRICAI 2008: Trends in Artificial Intelligence, vol 5351, 1104–1109
31. Wang T, Zhang X (2010) 3D protein folding structure prediction with genetic tabu search algorithm. Wuhan University of Technology, BMC Systems Biology
32. Kirbakaran P, Karthikeyan M, Singh KD et al (2012) In silico structural and functinal analysis of the human TOPK protein by structure modeling and molecular dynamics studies J Mol Model
33. Chira C, Horvath D, Dumitrescu D (2011) Hill-Climbing search and diversification within an evolutionary approach to protein structure prediction. BioData Mining 4
34. Rego C, Li H, Glover F (2011) A filter-and-fan approach to the 2D lattice model of the protein folding problem. Ann Operations Res 188:389–414
35. Junyan Z, Jinming L (2011) Application of simulated annealing algorithm in prediction of protein structure. Master Thesis of Fujian Agriculture and Forestry University