Ain Shams Engineering Journal 10 (2019) 713–719

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: www.sciencedirect.com

Electrical Engineering

Performance comparison of ab initio protein structure prediction methods

Mohamad Yousef [a,*], Tamer Abdelkader [a], Khaled El-Bahnasy [a]

[a] Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University, Abbassia 11566, Cairo, Egypt

ARTICLE INFO

Article history:
Received 18 September 2018
Revised 19 December 2018
Accepted 14 March 2019
Available online 6 April 2019

Keywords:
Genetic algorithm
Combinatorial optimization
Protein tertiary structure
Evolutionary algorithms
Ab initio prediction
De novo prediction
Free modeling

ABSTRACT

Wet Lab methods used to resolve the native structure of a given protein such as X-ray Diffraction and NMR spectroscopy are time-consuming, expensive, and could be done multiple times due to failure. Those pitfalls led to the development of computationally automated prediction methods. Predicting the structure of a protein using its peptide sequence only — also known as ab initio Protein Structure Prediction "PSP"— is computationally challenging because of the large conformational space to be searched and the complexity of energy functions. Some successful predictive methods have been developed to solve PSP problem. In this paper, by using a set of peptide sequences, we compare a collection of PSP ab initio-based methods. Experiments show that using a metaheuristic-based search method that utilizes genetic algorithm can achieve same or better results than time consuming methods.

1. Introduction

The function of any protein depends on its native 3D fold "structure" [1]. Proteins' native folds contribute in designing novel drugs, vaccines and proteins [2–4]. In the case where a protein fails to assume its correct fold it enters a state called misfolded protein which makes the protein fails to deliver its function. The formation of misfolded proteins leads to terminal forms of diseases such as: Alzheimer's, Parkinson's, diabetes, and some forms of cancer [5–8].

In vitro procedures used to discover proteins structures cannot minimize the difference between the number of protein sequences (≈120 M EBI (The European Bioinformatics Institute): as of 23-May-2018 [9]) and the number of sequences having solved structures (≈0.15 M PDB (Protein Data Bank): as of 29-May-2018 [10]). This is the main reason for developing PSP methods.

Any protein, which reaches its native state, should possess the lowest potential free energy. However, it is currently not possible to find a structure with the absolute global minimum energy for a given protein hence, the native correct structure. This is one of the fundamental setbacks in PSP. Furthermore, it is not yet computationally possible to discover all imaginable structures for a protein. It was predicted that a mainframe with a 10^{12} operations/second may work continuously for 10^{20} years to construct all imaginable structures of a peptide contains only 40 amino acids [11]. As a result, heuristic algorithms may become useful to reduce the huge search space.

Genetic Algorithm [12] (GA) is a metaheuristic algorithm based on the idea of natural selection and genetic-based operations such as crossover and mutation. John Holland presented it in 1975, then became widespread in the late 1980s [13] for its capability to solve different computational tasks. In the 1990s, Structural biology adopted GA to solve molecular docking problem [14]. In PSP field, multiple GA adaptations have been implemented, for example: HGA – a hybrid genetic algorithm for PSP which uses a path-relinking strategy [15] and A multiple minima genetic algorithm for PSP [16]. Later, we aim to focus on comparing some of PSP methods according to their output quality and the time needed to generate that output against our previously proposed method 3dProFold [29].

The upcoming sections are organized as follows: Section 2 presents a brief description about PSP and its basis. In Section 3 the PSP methods under study are reviewed. Section 4 discusses in detail the PSP method using GA. Section 5 presents the results of the experiments. Section 6 summarizes the conclusions.

* Corresponding author.
E-mail addresses: mohamad.yousef@cis.asu.edu.eg (M. Yousef), tammabde@cis.asu.edu.eg (T. Abdelkader), khaled.bahnasy@cis.asu.edu.eg (K. El-Bahnasy).
Peer review under responsibility of Ain Shams University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.asej.2019.03.004
2090-4479/

2. History and background

Levinthal's paradox in [17] states that: for a protein trying to assume its native structure, it is impossible for that protein to take all its imaginable conformational structures to finally reach its final fold due to the huge time needed for this process to occur. In vitro experiments showed that proteins take their final fold in seconds [18]. This indicates that every protein has a folding mechanism. Also, thermodynamic hypothesis [19] states that: the primary structure "sequence" of the polypeptide chain carries all the needed folding instructions. This also suggests that when the folding process ends, the final structure is restful and has reached its global minimum energy. Levinthal's paradox and thermodynamic hypothesis are the basis of ab initio PSP.

Also known as free modeling, ab initio prediction is preferred to be used if the target protein doesn't have a homologue - a homologue is a protein that is similar to the target protein by sequence - already existed in the biological databases. The task of predicting a target protein's structure would be relatively easy if a homologue was found and high-quality structure can be assembled by copying the framework of the homologue/s "solved" structure/s.

However, this procedure does not reveal any information about how and why a protein assumes its specific native structure. If no homologues are found, structures must be built from scratch. This defines the ab initio PSP concept.

Ab initio protein folding is considered a global optimization problem where the goal is to find the values of the dihedral angles for a given protein structure which contribute into that structure's stability and balance i.e.: having the global or near global minimum potential energy "Kcal/mol".

3. Related work

This section presents the methods under study. These methods were chosen for their promising prediction capability and their ease of use among the other ab initio-based methods.

3.1 QUARK

QUARK is a web server for ab initio PSP and protein peptide folding, it builds the correct protein structure using its amino acid sequence as an input. QUARK was the No. 1 server in Free modeling (FM) category in CASP9 and CASP10 contests [20,21]. QUARK is not a 100% ab initio method, because it uses a fragment assembly approach, in which tiny fragments (1–20 residues long taken from known PDB structures) are joined to build the final structure by replica-exchange Monte Carlo search with assistance of an atomic-level knowledge-based force field [22,23]. It accepts peptide sequences less than 200 amino acids long. In addition, QUARK users cannot submit multiple jobs at once.

3.2 PEP-FOLD2

Also known as PEP-FOLD 2.0, is an ab initio PSP web server aimed to carry out the process of PSP using amino acid sequences as an input. It is based on structural alphabet (SA) letters to construct the conformations of four consecutive residues, combines the predicted series of SA letters to a greedy algorithm and a coarse-grained force field. It may result in unconnected disulfide bonds in the all atom representation even if the cysteines are near each other. It accepts sequences ranged from 9 to 36 residues [24,25]. For an unknown reason, this method is prone to failure. It may require resubmitting the job multiple times.

3.3 PEP-FOLD3

Also known as PEP-FOLD 3.5, is a faster and more stable version of PEP-FOLD2. While it can deal with disulfide bonds, it cannot achieve results better than PEP-FOLD2. PEP-FOLD2 should be used for such peptides [26–28].

3.4 3dProFold

A desktop application that integrates GA search algorithm with ECEPPAK energy calculation program to find the best conformation of a protein using its amino acids sequence only which makes it a pure ab initio method. It does not use neither PDB templates nor fragment assembly [29]. This method is detailed in Section 4.

4. GA-based PSP

In this section, we are going to elaborate in depth about 3dProFold method, how GA phases are implemented, and how adaptations and integrations are made to make GA fitting to solve the PSP problem.

4.1 Encoding

Each individual in the population will be transformed into a 3D molecular representation using internal coordinates. These coordinates – also known as dihedral angles - have restricted values due to steric hindrance between the H atom in the NH group, the carbonyl oxygen, the H atom on the Cα and the side chain atoms located in the R group [30] - see Fig. 1. Each amino acid forms three common dihedral angles: Φ (Phi), Ψ (Psi), ω (Omega) which are the angles that define the mainchain. However, Amino acids differ in the number of the sidechain dihedral angles depending on their R group complexity. A single sidechain angle is called χ (chi) followed by a number indicating its succession in the R group: $\chi_1, \chi_2 \ldots \chi_i$. The first generation is created according to the Ramachandran plot to guarantee the quality of the generated individuals [31–34].

4.2 Selection

A subset of the population is chosen to generate offspring population using a selection scheme. Individuals will be chosen based on their fitness value measured by a fitness function (in our case an energy function). The fittest individuals will be a big part of the chosen individuals. Parents will be selected uniquely to prevent the creation of repeated individuals.

The majority of selection schemes are originated from a random selection approach. Most of them is implemented so that a small portion of the weak individuals would be chosen to help in keeping variety amongst the population. Tournament selection is one of the well-studied and common selection schemes.

Tournament selection scheme works by running several contests between the individuals. The winner of a contest is the individual with the smaller free/potential energy. Thus, it will be used as a parent in the upcoming step "crossover". "k" represents the selection intensity. As "k" increases, the tournament selection will prefer not to select weak individuals [35] which will eventually

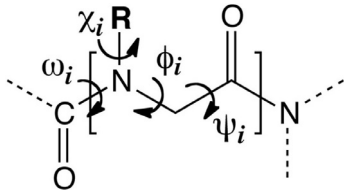

Fig. 1. Representation of the backbone and sidechain dihedral angles locations on the polypeptide chain atoms.

lead to the disappearance of the diversity from the generations [36]. According to our trials, k = 2 will suffice.

```
Procedure TournamentSelection (generation,k):
  SET optimal = 0.
  SET j = 1.
  while j < k.
    SET i= RandomNumber(1, N)
   SET ContenderSolution= generation[i]
if ((optimal == 0) || Calcfitness(ContenderSolution) better than Calcfitness (optimal))
then:
SET optimal = ContenderSolution.
End if.
  End while.
  return optimal.
End Procedure.
//N is the size of current generation.
```

4.3 Crossover

A couple of children will be formed by swapping the traits of the parents. The offspring individuals will share the traits of their parents. To avoid the early occurring of weak solutions and to maintain diversity across generations, we implemented 3 versions of crossover:

- 1-point crossover: Select one trait as a starting point "SP" and swap all the traits coming after.
- 2-point crossover:
- Select two traits and swap all traits between them.
- Uniform crossover:
- Swap every trait with crossover likelihood = 0.5.

4.4 Mutation

The GA permits for a minor chance of alteration to ensure that there will be no similar individuals generated by the crossover process. This is accomplished by going through all the traits of a single individual. Then, based on a mutation likelihood, a trait is selected to be changed, it can be altered by a minor change or replace it completely. The mutation likelihood usually ranges between 0.001 and 0.1. Mutation is incorporated to guarantee the diversity among the generations.

```
Procedure Mutation (chromosome,n):
    SET Chromosome2 = CLONE(chromosome)
    SET i = 1.
   while I < n.
      SET RndNumber= RandomFloat(0 to 1)
      if(RndNumber <= MR)
        SET Chromosome2[i] = Random(−180,180)
      End if.
End While.
Return Chromosome2.
End Procedure //Where MR is the predetermined mutation rate.
Procedure OnepointCrossover (parentA,parentB,n):
    SET SP= Random(2,n)
    SET j=SP.
    while j < n.
      SWAP (parentA[j],parentB[j])
    End while.
End Procedure.
//Where n is the length of a parent.
```

```
Procedure TwopointCrossover (parentA,parentB,n):
      SET SP1= Random(1,n/2)
      SET SP2 = Random(n/2,n)
      for k=SP1 to SP2.
        SWAP(parentA[k],parentB[k])
      End for.
End Procedure.
Procedure UniformCrossover (parentA,parentB,n):
  SET y = 1.
  while y < n.
    SET RndProb= RandomFloat(0 to 1)
    if(RndProb <= 0.5)
      SWAP(parentA[y],parentB[y])
End if.
  End while.
End Procedure.
```

4.5 Structural Enhancement

The crossover process swaps angle values between individuals regardless of the offspring's structural stability. Protein structure stability is measured by its free energy. High free energy values are a signal of a disrupted structure. A refinement step is required to enhance the stability of an individual. This is done by integrating ECEPPAK feature with the GA workflow [37] – see Fig. 2.

4.6 Elitisim

In each generation, the fittest individual will be added to an ultimate list of candidate final structures. Those will be superimposed to the original structure and the candidate with the maximum similarity to the original structure will be selected as the final solution.

4.7 Fitness calculation

Each individual will have its fitness calculated by the following force field equation implemented by ECEPPAK:

$$E_{tot} = E_{es} + E_{nb} + E_{tor} + E_{loop} \quad (1)$$

where

- E_{tot}: Total energy (fitness value).
- E_{es}: Energy of Electrostatic interactions.
- E_{nb}: Energy of Nonbonded interactions.
- E_{tor}: Sidechain torsional energy.
- E_{loop}: Energy of Loops interactions.

Both Ramachandran plot and 3 crossover versions were implemented to cover the absence of:

1. Surface hydration model – with solvent and atomic hydration parameters that are optimized using nonpeptide data (the ECEPPAK uses solvation parameters file "srfopt.set" for this feature)
2. Volume hydration model (the ECEPPAK uses solvation parameters file "volume.set" for this feature).
3. EDMC "Electrostatically Driven Monte Carlo" calculations which has shown its consumption for a lot of time, nearly a week to 2 weeks to predict a small peptide chain of 31 residues "1Q2K", and its low quality produced conformations. Also, the work in [38] showed that EDMC calculations option in ECEPPAK produced structures with low quality compared to native.

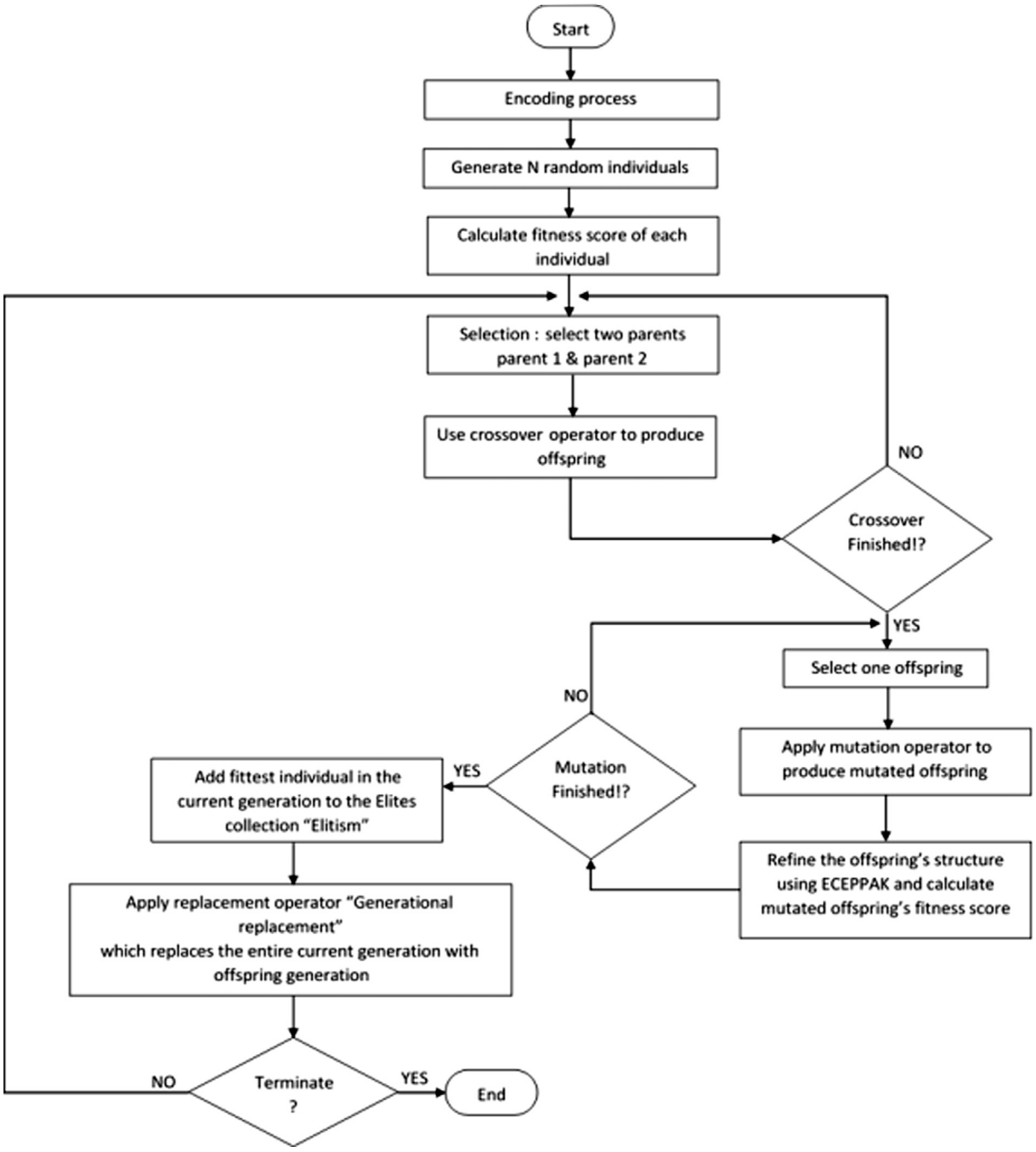

Fig. 2. GA overall framework after integration with the energy minimization feature.

5. Experimental results

Using the following PDB entries: 1Q2K [39], 1DU9 [40], 1BH0 [41] and 1WM7 [42] of lengths: 31, 28, 29, 29 residues respectively, multiple runs have been performed to assess the accuracy of the predictive ab initio-based computational methods: QUARK, PEP-Fold2, PEP-Fold3 and our previous work 3dProFold. Then, the produced structures were compared to the original structures using TM-Align [43].

TM-Score and RMSD are the two main criteria that TMAlign uses to measure the similarity between 2 input structures and they are also approved by the CASP community. If the TM-Align reported similarity measures of $0.5 \leq$ TM-Score ≤ 1 and RMSD < $\approx$3 Å (1 Å = 10^{-10} m) then, the target structure is around the same fold as the native one.

From Tables 1 and 3 (and also Figs. 3 and 5), the TM-Scores shows that the 3dProFold method surpassed the other methods in predicting the structures of 1BH0 and 1Q2K but took a longer time compared to them. In structure 1DU9, PEP-FOLD2 has predicted its structure in 0.5 hrs with a slight difference in TM-Score ahead of 3dProFold and QUARK which took very long to predict its structure "14 hrs". 1WM7 structure prediction results shows that 3dProFold is as near as QUARK in terms of structural similarity and it took 3dProFold shorter time compared to QUARK. In Table 2 – also refer to Fig. 4 – all the predicted structures have RMSDs values $\approx$3 Å which is an additional indicator of good accuracy shared among the methods.

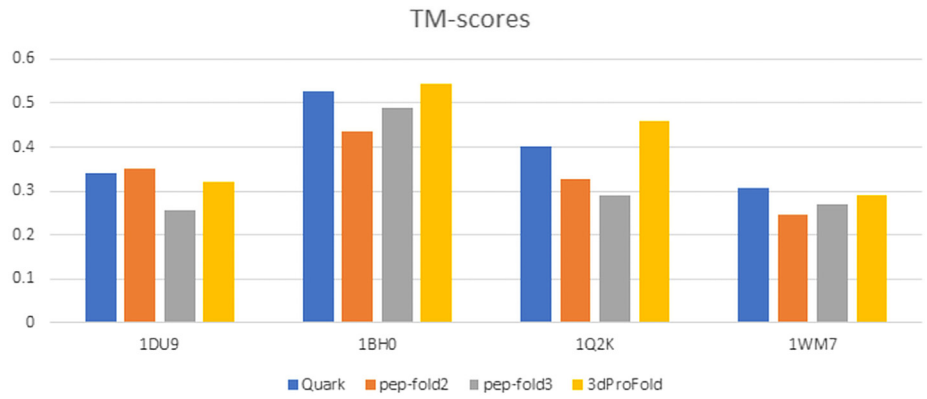

Fig. 3. TM-Scores (the greater the better).

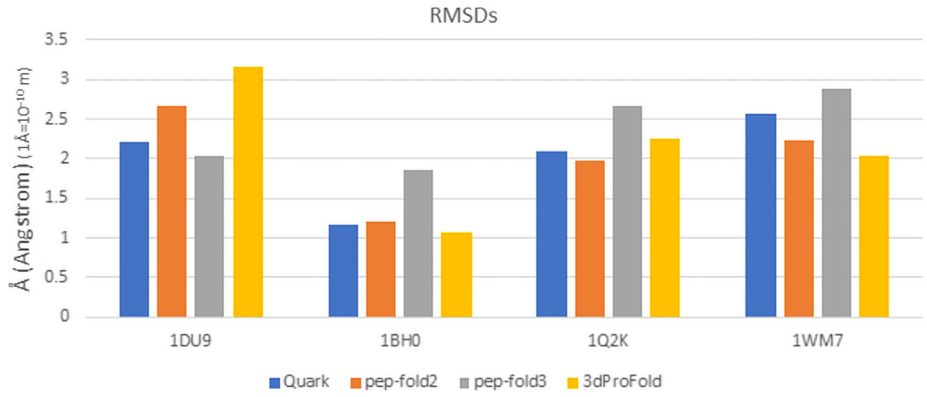

Fig. 4. RMSD values (lesser than 3 Å is better).

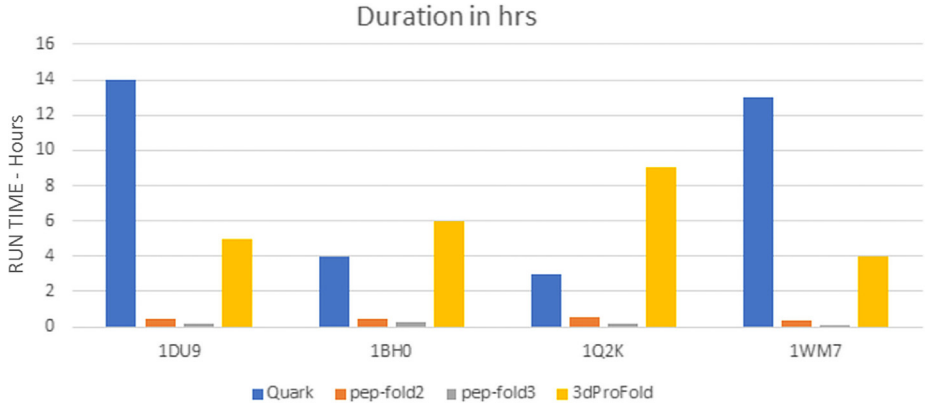

Fig. 5. Time consumed – in hrs.

6. Conclusion and discussion

In our work we compared a collection of Protein Structure Prediction "PSP" methods against our implemented method 3dProFold in terms of their accuracy and time consumption. We aimed to find the best performing method in the context. The time consumption of PEP-FOLD2 and PEP-FOLD3 is much smaller than the other methods but their accuracies are not guaranteed to be

Table 1
TM-Scores of structures obtained.

PDB_ID	QUARK	Pep-Fold2	Pep-Fold3	3dProFold
1DU9	0.34	0.35	0.256	0.32
1BH0	0.527	0.435	0.49	0.545
1Q2K	0.4	0.326	0.29	0.46
1WM7	0.306	0.245	0.27	0.289

Table 2
RMSDs of structures obtained.

PDB_ID	QUARK	Pep-Fold2	Pep-Fold3	3dProFold
1DU9	2.22	2.67	2.04	3.17
1BH0	1.17	1.21	1.86	1.07
1Q2K	2.1	1.98	2.67	2.25
1WM7	2.57	2.23	2.89	2.04

Table 3
Time consumption "in terms of Hrs"

PDB_ID	QUARK	Pep-Fold2	Pep-Fold3	3dProFold
1DU9	14	0.5	0.2	5
1BH0	4	0.5	0.25	6
1Q2K	3	0.53	0.2	9
1WM7	13	0.36	0.1	4

high. Both QUARK and 3dProFold may forsake time in favor of accuracy. The long time consumed by 3dProFold method is attributed to the overhead of Ramachandran plot feature implemented.

In 3dProFold method, GA is implemented to search a compact conformational space of refined structures in attempt to find optimal/near optimal structure using a given peptide's amino acid sequence only as an input (De novo/Ab initio prediction), because of the great potential GA has and its ability to model and solve a variety of problems. However, the generated structures may be distorted and noisy. An enhancement step was added to confirm that the protein structures are at rest in the 3D space.

The absence of surface hydration model could be the reason behind the setback in 3dProFold resulted structures. Its absence is attributed to the energy function implemented in ECEPPAK which takes a longer time when using surface hydration model without any improvement noticed in the results and in case of volume hydration model the ECEPPAK crashes and aborts due to errors found in the volume solvation parameters file "volume.set".

Saving time should speed up the process of discovering novel drugs and biological vaccines, which will benefit both the drug industry and patients waiting for new treatments. Our future work will be focused on enhancing performance and reducing time consumption using CUDA GPU technology.

References

[1] Alberts B, Johnson A, Lewis J, et al. Molecular biology of the cell. 4th ed. New York: Garland Science; 2002. The shape and structure of proteins. Veerapandian P, editor, Structure-based drug design. New York: Marcel Dekker, 1997.
[2] Breda A, Valadares NF, De Souza ON, Garratt RC. Prot Struct, Model Appl 2008.
[3] Lilie H. Designer proteins in biotechnology. EMBO Rep 2006;4(4):346–51.
[4] Jaenicke R, Sterner R. Protein design at the crossroads of biotechnology, chemistry, theory, and evolution. Angew Chem Int Ed 2003;42(2):140–2.
[5] Smith Adam. Protein misfolding. Nat Rev Drug Discov 2003;426 (6968):78–102.
[6] Dobson CM. Protein folding and misfolding. Nature 2003;426(6968):884–90.
[7] Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem 2006;75(1):333–66.
[8] Mathias Jucker, Walker Lary C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. Nature 2013;501(7465):45–51.
[9] <http://www.ebi.ac.uk/uniprot/TrEMBLstats>.
[10] <https://www.wwpdb.org/stats/deposition>.
[11] Xiong J. Essential bioinformatics. Cambridge University Press; 2006.
[12] Holland JH. Adaptation in natural and artificial systems. Ann Harbor, MI: The University of Michigan Press; 1975.
[13] Goldberg DH. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison-Wesley; 1985.
[14] Willett P. Genetic algorithms in molecular recognition and design. Trends Biotechnol 1995;13(12):516–21. doi: https://doi.org/10.1016/S0167-7799(00)89015-0. PMID 8595137.
[15] Dorn M, Buriol LS, Lamb LC. A hybrid genetic algorithm for the 3-d protein structure prediction problem using a path-relinking strategy. In: Proceedings of the IEEE congress on evolutionary computation. p. 2709–16.
[16] Custódio FL, Barbosa HJC, Dardenne LE. A multiple minima genetic algorithm for protein structure prediction. Appl Soft Comput 2014;15:88–99.
[17] Levinthal C. Are there pathways for protein folding? J Med Phys 1968;65 (1):44–5.
[18] Hartl F Ulrich, Hayer-Hartl Manajit. Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol 2009. doi: https://doi.org/10.1038/nsmb.1591.
[19] Anfinsen CB. The principles that govern the folding of protein chains. Science 1973;181(4096):223–30.
[20] <http://predictioncenter.org/casp9/CD/data/html/groups.server.fm.html>.
[21] <http://predictioncenter.org/casp10/groups_analysis.cgi?type=server&tbm=on&tbm_hard=on&tbmfm=on&fm=on&submit=Filter>.
[22] Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 2012;80:1715–35.
[23] Xu D, Zhang Y. Toward optimal fragment generations for ab initio protein structure assembly. Proteins 2013;81:229–39.
[24] Maupetit J, Derreumaux P, Tufféry P. PEP-FOLD: an online resource for de novo peptide structure prediction. Nucl Acids Res 2009;37. doi: https://doi.org/10.1093/nar/gkp323.
[25] Maupetit J, Derreumaux P, Tuffery P. A fast and accurate method for large-scale de novo peptide structure prediction. J Comput Chem. 2010. 31–726-38.
[26] Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. Nucl Acids Res 2016;8(44):W449–54.
[27] Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J Chem Theor Comput 2014;10:4745–58.
[28] Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucl Acids Res 2012;40:W288–93.
[29] Yousef M, Abdelkader T, ElBahnasy K. A hybrid model to predict proteins tertiary structure. In: 2017 12th International conference on computer engineering and systems (ICCES), Cairo; 2017. p. 85–91. 10.1109/ICCES.2017.8275282.
[30] Liljas Anders, Liljas Lars, Piskur Jure, Lindblom Göran, Nissen Poul, Kjeldgaard Morten. Textbook of structural biology. World Scientific; 2009
[31] Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol 1963;7:95–9. doi: https://doi.org/10.1016/S0022-2836(63)80023-6. PMID 13990617.
[32] Kleywegt GJ, Jones TA. Phi/psi-chology: Ramachandran revisited. Structure 1996;4(12):1395–400. doi: https://doi.org/10.1016/S0969-2126(96)00147-5.
[33] Hovmöller S, Zhou T, Ohlson T. Conformations of amino acids in proteins. Acta Crystallogr D 2002;58(Pt 5):768–76. doi: https://doi.org/10.1107/S0907444902003359.
[34] Anderson RJ, Weng Z, Campbell RK, Jiang X. Main-chain conformational tendencies of amino acids. Proteins 2005;60(4):679–89. doi: https://doi.org/10.1002/prot.20530.
[35] Miller Brad, Goldberg David. Genetic algorithms, tournament selection, and the effects of noise. Complex Syst 1995;9:193–212.
[36] Blickle Tobias, Thiele Lothar. A comparison of selection schemes used in evolutionary algorithms. Evol Comput 1996;4(4):361–94. doi: https://doi.org/10.1162/evco.1996.4.4.361.
[37] Scheraga HA, Ripoll DR, Liwo A, Czaplewski C. User Guide ECEPPAK and ANALYZE Programs.
[38] Barrera Guisasola EE, Masman MF, Enriz RD. cent.eur.j.chem. 2010;8:566. doi: https://doi.org/10.2478/s11532-010-0015-1.
[39] Cai Z, Xu C, Xu Y, Lu W, Chi CW, Shi Y, et al. Solution structure of BmBKTx1, a new BKCa1 channel blocker from the Chinese scorpion Buthus martensi Karsch. Biochemistry 2004;43:3764–71.
[40] Xu Y, Wu J, Pei J, Shi Y, Ji Y, Tong Q. Solution structure of BmP02, a new potassium channel blocker from the venom of the Chinese scorpion Buthus martensi Karsch. Biochemistry 2000;39:13669–75.
[41] Sturm NS, Lin Y, Burley SK, Krstenansky JL, Ahn JM, Azizeh BY. Structure-function studies on positions 17, 18, and 21 replacement analogues of glucagon: the importance of charged residues and salt bridges in glucagon biological activity. J Med Chem 1998;41:2693–700.
[42] Wu G, Li Y, Wei D, He F, Jiang S, Hu G, et al. Solution structure of BmP01 from the venom of scorpion Buthus martensii Karsch. Biochem Biophys Res Commun 2000;276:1148–54.
[43] Zhang Y, Skolnick J. TM-align: A protein structure alignment algorithm based on TM-score. Nucleic Acids Res 2005;33:2302–9.

Mohamad Yousef received a B.Sc. degree in bioinformatics in 2013 from Ain Shams University, Cairo, Egypt. He worked in several Software warehousing companies. Currently, he is a teaching assistant at Ain Shams University – FCIS. His current research interests include Proteins structure prediction, Genetic Algorithm, Evolutionary Algorithms, Protein Free Modelling, Artificial intelligence, Data Science, bioinformatics, computational biology, and applied soft computing.

Dr. Tamer Abdelkader received a B.Sc. degree in electrical and computer engineering and an M.Sc. degree in 2003 from Ain Shams University, Cairo, Egypt. He received M.Sc. and Ph.D. degrees in 2012 in electrical and computer engineering from the University of Waterloo, Ontario, Canada. He worked in the University of Waterloo as a postdoctoral, and a network consultant in Egypt. Currently, he is an assistant professor at Ain Shams University. He is the author of several publications in IEEE journals and conferences. His current research interests include mobile computing, vehicular networks, energy-efficient protocols, mobile social networks, and applied soft computing.