Malaysian Journal of Fundamental & Applied Sciences

available online at http://mjfas.ibnusina.utm.my

In Silico Protein Folding Simulation of Amyloid A4 Peptide

Fatahiya Mohamed Tap[1] and Nurul Bahiyah Ahmad Khairudin[2*]

[1]Department of Bioprocess, Faculty of Chemical Engineering, UTM, 81310 UTM Skudai, Johor, Malaysia
[2]Malaysia Japan International Institute of Technology, UTM International Campus, 54100 Kuala Lumpur, Malaysia

ABSTRACT

The objective of this study is to investigate the folding pathways of Amyloid A4 peptide (PDB ID: 1AML). The structure and trajectories of this protein has been studied using Molecular Dynamics (MD) simulation. The simulation was run at 300K for 50nsec. The model at 2nsec was aligned to the Nuclear Magnetic Resonance (NMR) structure with the RMSD value of 7.85Å for the overall structure. It was found that the 3_10_helix started to form (ILE32 GLY33 LEU34 MET35) at 4nsec until 7nsec.

| Amyloid A4 peptide | Molecular Dynamics | Protein folding |

1. INTRODUCTION

Understanding protein folding process is important in molecular biology field because several diseases such as Alzheimer and cancer are directly related to the misfolding of protein. All these diseases have no cure until today and this problem has not been solved for more than 40 decades.

Some researchers claimed that, the cure of those diseases can be found when folding process of protein is known. Unfortunately, this is the major challenge in molecular biology today since nobody knows how the protein folds. Protein folding is a process in which, the sequence of amino acids fold naturally into its tertiary structure. The Formation of the protein tertiary structure is related to the interactions among the amino acid residues [1]. For example, the most important finding regarding on understanding of protein folding was carried out by Anfinsen and his colleagues in which, they claimed that the structure of the protein is solely determine by the sequence of amino acids [2]. Findings by Anfinsen and colleagues have inspired researchers to continue investigating the pathways of protein folding.

Since the evolution in studying protein folding increased rapidly, researchers have come up with various methods and proven that protein folding can be simulated using computers [3]. Computational method is the most powerful tool because it is extremely high resolutions and can provide detailed atomic level presentations. Furthermore, the increase in computer speed and improvements in force field along with more efficient computation algorithms have brought realistic computer simulation of the folding process

within reach. Several works have been carried out using this method [4-5].

In this study, implicit solvent method was employed to study the 1AML protein folding. This protein is a small peptide present in normal cells and aggregated 1AML is the main constituent of the extracellular amyloid plaques found in Alzheimer's disease [6]. The main information about Amyloid, it is actually not toxic [7], however it may become toxic to neuronal cells once it has aggregated into amyloid fibrils by peptide-peptide interaction. Thus, this study aims to investigate further on the early stage of the folding in order to contribute better understanding of the folding pathways of this protein.

2. EXPERIMENTAL

2.1 Material and method

The Protein used in this study was Amyloid A4 peptide [6]. From the previous research this protein was a deduction from β-amyloid which is a big protein that was found in human brain. This protein contains 40 residues and chain A. This protein was simulated using an all atom classical simulation and slightly modified version of AMBER 11[8] force field 99 [9]. The energy minimization was carried out using 1000 cycles of steepest descent and another 1000 cycles of conjugate gradient [10]. The MD simulation was carried out for 50ns. The solvation effects were incorporated with Generalized Bond implicit solvent method [11]. The system used Langevin dynamics [12] with the collision frequency is 1.0 to maintain the temperature at 300K. Bond constraints

*Corresponding author at:
E-mail addresses: nurulbahiyah@ic.utm.my (Nurul Bahiyah)

were imposed on a ll bonds involving hydrogen atoms via SHAKE [13].

3. RESULTS & DISCUSSION

3.1 Collapse and Contacts Occurred from 1nsec to 50nsec.

The unfolded starting structure was at extended state with very few native contacts and no helical content. Most of the native contacts were absent at the unfolded state. The trajectories of this protein were different from 1nsec to 20nsec, after 20nsec until 50nsec the structures looked similar and more stable. During the simulations, from linear chain the first collapse occurred at 2nsec in which GLY33 contacts with

ILE31 with distance 6.64Å. F igure 1.0 showed the residue contacts formed in the model.

3.2 RMSD

The $RMSD_{back-all}$ was calculated to be 10Å at 50nsec as compared to that of the NMR structure. Figure 2.0 shows the $RMSD_{back-all}$ from 1nsec to 50nsec. At the starting point of the simulation, RMSD value is 30.25Å. However, the RMSD value decreased as the simulation progressed. H igh fluctuations were occurred from 0nsec to 5nsec. T he RMSD values somehow look more stable with stable fluctuation from value 10.0Å to 5.0Å within 10nsec to 50nsec. It was observed that the model showed the lowest $RMSD_{back-all}$ value 7.85Å at 2nsec. Figure 3.0 shows the superimposed of the model at 2nsec with that of the native structure.
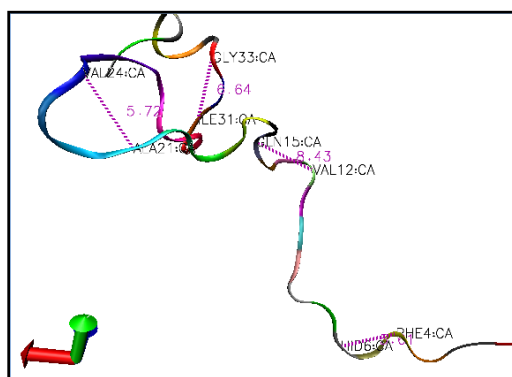

Fig. 1 Contacts between ALA24–ALA21, GLY33–ILE31, GLN15-VAL12 and PHE4-HID6
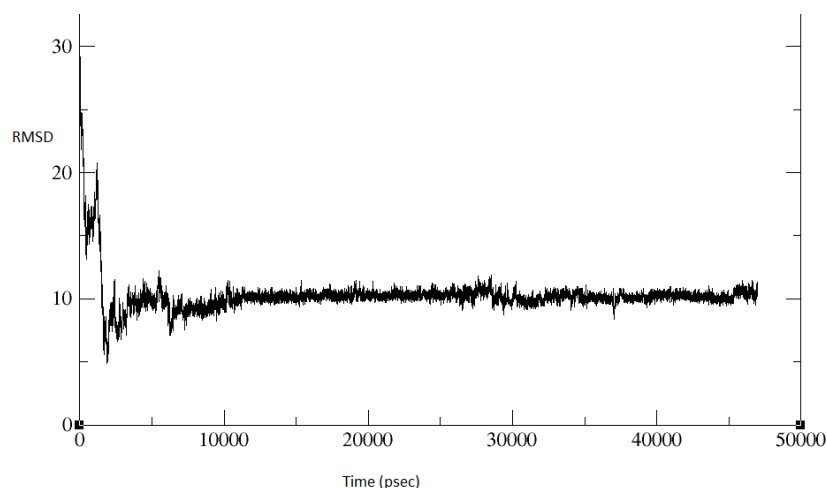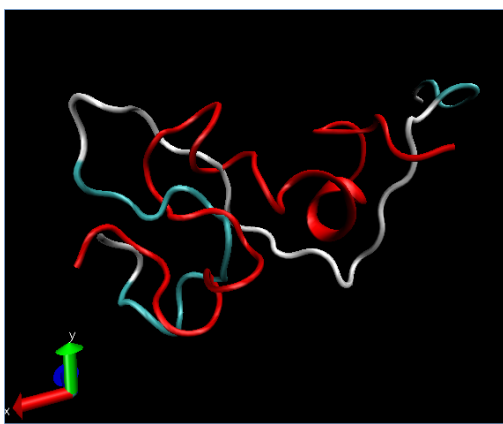

**Fig. 2** $RMSD_{back-all}$ value within from 1nsec to 50nsec

Fig. 3 Superimposed of the model at 2 ns with native NMR structure. The Red structure represents native structure and the other structure represents model structure.
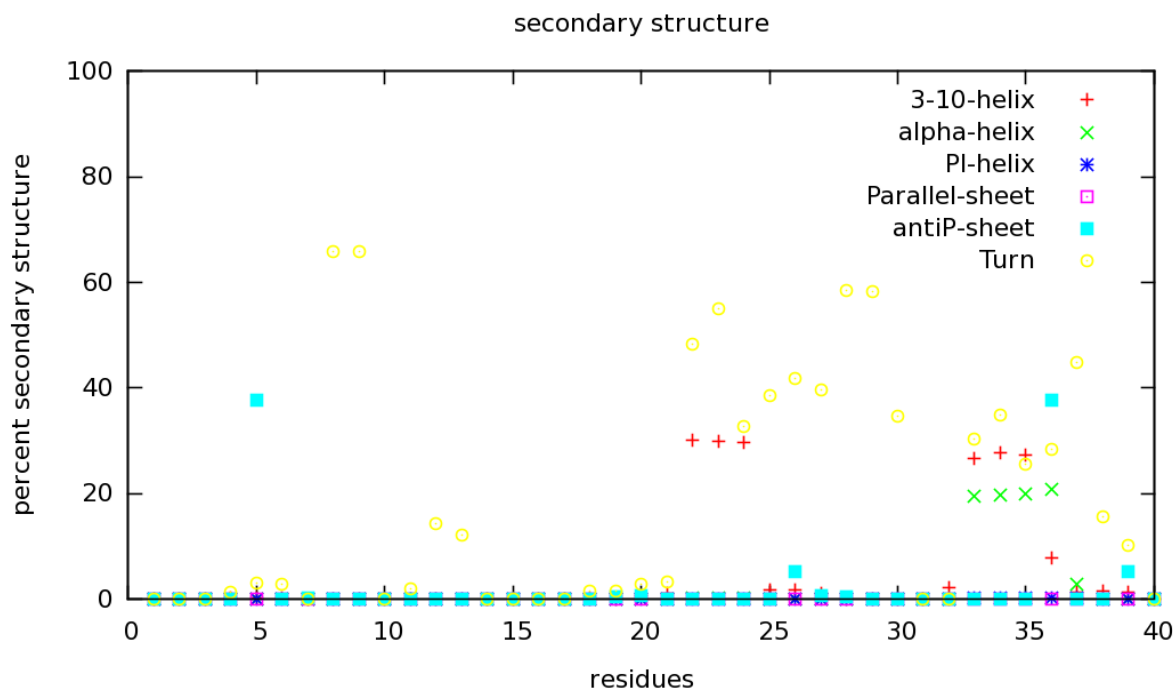
3.3 Secondary structure

Previous studies have shown that secondary structures will form from nanosecond to microseconds. For the model, it was found that the 3_{10}-helix started to form (ILE32 GLY33 LEU34 MET35) at 4nsec until 7nsec. Figure 4.0 shows the structure of 3_{10}-helix at 7nsec. The NMR structure of this protein contains helices at (ILE31 ILE32 GLY33 LEU34 MET35) and (GLN15

LYS16 LEU17 VAL18 PHE19 PHE20 ALA21 GLU22 ASP23). Previous findings from experimental method had shown that helix formation formed from Val12 to Met35 [7]. This suggested that the formation of

3_{10}-helix will lead to the formation of helix (ILE31 ILE32 GLY33 LEU34 MET35) when the contacts among residues or hydrogen bonds become more stable and this protein should be further simulated. The graph at Figure 5.0 shows the percentage of secondary structures formed for each residue within 1nsec to 50nsec. The data were gathered using the ptraj analysis. From the graph, it can be observed that the formation of 3_{10}-helix was formed between residues GLU22 to GLY25 and GLY33 to MET35 with the percentage between 20% - 40% for 50nsec. There was no occurrence of β-sheets observed throughout the whole simulation.
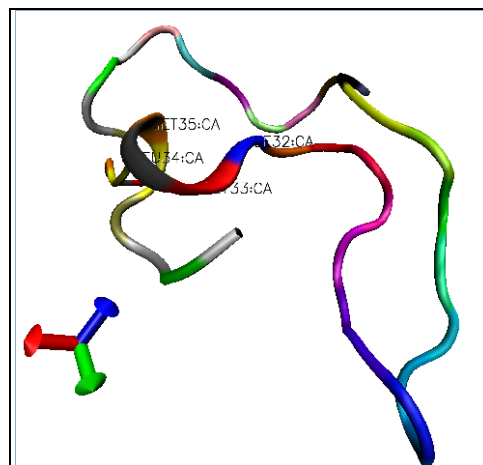
Fig. 4 Structure of 3_{10}-helix at 7nsec

Fig. 5 The Percentage of Occurrence of secondary structures for each residue from 1nsec to 50nsec.

4.   CONCLUSION

In this study, MD was used to study the folding pathway of Amyloid A4 peptide with simulation time of 50nsec. The first collapsed occurred at 2nsec and produced contacts between ALA24–ALA21, GLY33–ILE31, GLN15-VAL12 and PHE4-HID6 with the distance less than 10Å. The lowest RMSD value (7.85Å) for this model was at 2nsec. The 3_{10}-helix was formed between residues GLU22 to GLY25 and GLY33 to MET35 and no occurrence of β-sheets for 50nsec simulation.

ACKNOWLEDGEMENT

REFERENCES

[1]   A.R. Leach, Molecular Modelling: Principle and application, 2nd edition. Pearson Education, 2001.

[2]   E. L. Kavraki., Protein folding, unpublished.

[3]   D.Bratko, T. Cellemer, J.M. Prausnitz, and H.W. Blanch., Biotechnology and Bioengineering, 9 2007,1-8.

[4]   C.B. Anfinsen, Science, 181 (1973), 223-230.

[5]   N. Ferguson, and A.R. Fersht, Elsevier,13 (2003), 75-81.

[6]   H. Stricht, P. Bayer, D. Willbold, S. Dames, C. Hilbich, K. Beyreuther, W.R.Frank, and P. Rosch , Eur. J. Biochem, 233 (1995), 293-298.

[7]   K. Wataba, T. Segawa, K. Nakamura, M. Kodaka, T. Konakahara, and H. Okuni, J. Peptide Res., 58 (2001), 342-346.

[8]   D.A Case, T.A Darden, T.E. Cheatham, III, C.L Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A Kollman, Amber 11, University of California, San Francisco, 2010.

[9]   D.A. Case, T. cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr.,A. Onufriev, C. Simmerling, B. Wang and R. Woods, J. Computat. Chem, 26 (2005), 1668-1688.

[10]   V. Tsui, and A.D. Case, Biopolymers (Nucleic Acid Sciences), 56 (2001), 275-291.

[11]   D.J. Sindhikara, S. Kim, A.F. Voter, and A.E Roitberg, J. chem.. Theory. Comput.,5 (2009), 1624-1631.

[12]   J.P. Ryckppaert, G. Ciccotti, and H.J.C Berendsen, J. Comput. Phys., 23 (1977), 327-341.