STRUCTURAL ANALYSIS AND MODELING COMPARISON OF PRIMATES' AMYLOID βA4 PROTEIN

FAHIMEH JALALI

A dissertation submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Biotechnology)

Faculty of Bioscience and Bioengineering
University Technology Malaysia

January 2013

To my beloved family and fiancé

ACKNOWLEDGMENT

I would like to begin by sincerely thanking my supervisor, Associated Prof Dr. Shahir Shamsir for his constant support, guidance and mentorship over the course of this thesis. He gave me the freedom to define my thesis statements. I am very grateful to having the opportunity for work on this thesis.

I am heartily thankful to my parents, for their unconditional supports and encouraging me to continue my education. I am very grateful to my lovely fiance for his professional gaudiness and emotional supports.

ABSTRACT

Amyloid protein originally termed "beta-protein" or "amyloid A4" which is indicated as "beta A4." Amyloid beta A4 protein is proteolytically derived from a transmembrane protein named amyloid precursor protein (APP) which is encoded by an extensively expressed gene on chromosome 21. Mutations in Amyloid βA4 gene cause the plaques which are composed of a tangle of regularly ordered fibrillar aggregates called amyloid fibers, and abnormal accumulation of amyloid fibrils make protein misfolding diseases so that it leads to amyloidosis and neurodegenerative disorder like Alzheimer diseases (AD) or Parkinson diseases (PD).Currently there are 91 structure of Amyloid beta A4 protein have been predicted in various species, but there is no presently concrete evidence that shows Amyloid beta A4 model in primates. This study was focus on Amyloid beta A4 protein information and analyzed using bioinformatics software such as: Uniprot, Deep View, PDB (protein data bank), Swiss Model, Jal View, BLAST, VMD(Visual molecular dynamics), NCBI(National center for biotechnology information) therefore those of programs help us to predict and create a new model of 3D structure for this protein and they are useful for analysis multiple alignment, simulation , image processing for βA4 protein sequence, illustrate conserved regions and residues of protein between different species (human being and primates) to indicate comparison of their structural features and their gen properties to this end 8 primates were chosen. Based on the analysis of this comparison demonstrated that some of the primates are highly conserved with their template (Homo sapiens) and they have similar primary and tertiary structure with template .The reason of this issue is that in all of them the protein's gene location is on chromosome 21 as same as human gen location. On the other hand rest of chosen primates is not conserved with template and their structures are totally different with Homo sapiens due to APP gene location is on chromosome 3 instead of chromosome 21. This information was gathered from GenBank which is the genetic sequence database, an annotated collection of all publicly available DNA sequences.

ABSTRAK

Amiloid protein asalnya dipanggil "beta-protein" atau "A4 amyloid" yang ditandakan sebagai "A4 beta." Beta amiloid protein A4 proteolytically diperolehi daripada protein transmembran pelopor protein amiloid dinamakan (APP) yang dikodkan oleh gene meluas dinyatakan pada 21 kromosom. Mutasi dalam gen βA4 amiloid menyebabkan plak yang terdiri daripada kekusutan agregat kerap diperintahkan berhubung dgn urat saraf dipanggil serat amiloid, dan pengumpulan abnormal gentian halus amiloid membuat penyakit misfolding protein supaya ia membawa kepada amyloidosis dan gangguan neurodegenerative seperti penyakit Alzheimer (AD) atau Parkinson penyakit (PD). Pada masa ini terdapat 91 struktur protein amiloid beta A4 telah diramalkan dalam pelbagai spesis, tetapi tidak ada bukti yang kini konkrit yang menunjukkan amiloid beta A4 model dalam primat. Kajian ini adalah memberi tumpuan kepada protein amiloid beta maklumat A4 dan dianalisis menggunakan perisian bioinformatik seperti: Uniprot, Deep View, PDB (protein data bank), Model Switzerland, Jal View, letupan, VMD (Visual molekul dinamik), NCBI (National pusat maklumat bioteknologi) itu mereka program membantu kita untuk meramalkan dan mewujudkan satu model baru struktur 3D protein ini dan mereka adalah berguna untuk penjajaran pelbagai analisis, simulasi, pemprosesan imej untuk βA4 urutan protein, menggambarkan kawasan terpelihara dan sisa protein antara spesies yang berbeza (manusia dan primat) untuk menunjukkan perbandingan ciri-ciri struktur mereka dan sifat-sifat gen mereka untuk tujuan ini 8 primat telah dipilih. Berdasarkan analisis perbandingan ini menunjukkan bahawa beberapa primat sangat dipelihara dengan template (Homo sapiens) mereka dan mereka mempunyai struktur yang serupa utama dan tertiari dengan template. Sebab isu ini adalah bahawa dalam semua daripada mereka lokasi gen-protein ini adalah pada 21 kromosom yang sama sebagai lokasi gen manusia. Pada rehat tangan lain primat yang dipilih tidak dipelihara dengan template dan struktur mereka adalah sama sekali berbeza dengan Homo sapiens disebabkan lokasi gen APP adalah pada 3 kromosom bukannya kromosom 21. Maklumat ini dikumpulkan dari GenBank yang merupakan pangkalan data jujukan genetik, koleksi beranotasi semua jujukan DNA yang boleh didapati oleh orang ramai.

TABLE OF CONTENTS

4 RESULTS AND DISCUSSION

5 CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AD	Alzheimer Diseases
PD	Parkinson Diseases
Aβ	Amyloid Beta
SNCA	Alpha_synuclein gen
GDNF	Glial Derived Neurotrophic Factor
GAD	Glutamic Acid Decarboxylase
APP	Amyloid Precursor Protein
UniProt	Universal Protein resource
PDB	Protein Data Bank
NCBI	National Center for Biotechnology Information
DS	Down's Syndrome
HPC	High Pathology Control
APOE	Apolipo Protein E
3D	Three Dimensional
MD	Molecular dynamic
SPC	Simple Point Charge
PBC	Periodic Boundary Conditions
PME	Particle Mesh Ewald
RMSD	Root Mean-Square Deviation
RACE	Rapid Amplification CDNA Ends
RMSF	Root Mean Square Fluctuations
MSA	View and edit Multiple Sequence Alignment
VMD	Visual Molecular Dynamic
AcD	Acidic Domain
GFLD	Growth Factor-Like Domain
CuBD	Copper-Binding Domain

AAV_GDNF	Adeno_Associated Viral vector _ Glial Derived Neurotrophic Factor
GABA	Gamma-Aminobutyric Acid
BLAST	Basic Local Alignment Search Tool
ND	Neurodegenerative Disease
HMM	Hidden Markor Models
GROMACS	Groningen Machine for Chemical Simulation
NMR	Nuclear Magnetic Resonance spectroscopy
QMEAN4	Qualitative Model Energy Analysis
RACE	Rapid Amplification of CDNA

LIST OF SYMBOLS

DNA	Deoxyribonucleic Acid
RNA	Reoxyribonucleic Acid
CHI	Chitinase
Ca	Calcium
cDNA	Complementary DNA
G. antarctica PI12	Glaciozyma antarctica PI12
MPTP	Methyl_Phenyl_Tetrahydro Pyridine
K^0.	Kelvin
GLU	Glutamic acid
Asp	Aspartic acid
Arg	Argenine
A°	Angstrom
aa	amino acid
Z score	Standard score
TM score	Template modelling score
Gly	Glycine
Phe	Phenylalanin
Ile	Isoleucine
Ala	Alanine
Thr	Theronine
His	Histidine
Lys	Lysine
H_bond	hydrogen bond
NE	Noradrenaline neuron
1D2k	The Xray structure of a chitinase from the Pathogenic fungus Coccidioides

NH	Neuron
3NYL	PDB cod of The X-ray structure of an antiparallel dimer of the humaamyloid precursor protein E2 domain
°C	Celsius

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.0 Background of Studies

Protein can be defined as biochemical compounds which are consisting of one or more polypeptide generally folded into globular or fibrous form. Some of them are insoluble fibrous protein named amyloid so that they increase from minimum eighteen unsuitable folded proteins and polypeptides that exist normally in the body. The misfolded structure of these proteins and polypeptides have been incorporated with the pathology of more than twenty severe human illnesses in that unnatural reposition of amyloid fibrils in organ might cause amyloidosis and neurodegenerative disorder, like Alzheimer diseases (AD) and Parkinson diseases (PD).

1.0.1 Description of protein

There are different type of amyloid and amyloid beta or (Aβ) is one of significant ones that is the most important portion of amyloid plaques which detected in the brain of patients with AD, in addition same plaques emerge in some versions of lewy body dementia and inclusion body myositis (muscle disease).

The plaques are contained of a tangle of orderly arranged fibrillar accumulation named amyloid fibers, a protein folded partake by other peptides like the prions related within protein misfolding diseases, Tha soluble oligomeric shaped of the peptide might be causative factor in the progress of AD diseases.

1.0.2 Disease and exclusive mutations

As mentioned Alzheimer's and Parkinson's disease are associated with the formation in the brain of amyloid fibrils from β_amyloid and α_synuclein proteins respectively. Alpha_synuclein is a protein in human is encoded by SNCA gene, usually found in an amyloid fraction is shown to be a fragment of its precursor protein which cause Alzheimer diseases amyloid.

Actually amyloid's mutants such as: (A53T,A3OP,E22G) coupled with familial Alzheimer's and Parkinson's diseases from morphologically identical annular protofibrils that look like a class of pore_forming bacterial toxin, suggesting that inappropriate

membrane permeabilization might be the cause of cell dysfunction and even cell death in amyloid dieases.

According to Lang,A.E & Lozano,A.M (1998) Parkinson disease is a common neurodegenerative disorder characterized by the loss of dopaminergic neurons and the presence of intracytoplasmic-ubiquitinated inclusions (Lewy bodies). Mutations in α-synuclein (A53T, A30P) and parkin cause familial Parkinson disease. Both these proteins are found in Lewy bodies.(Lewy bodies are abnormal aggregates of protein that develop inside nerve cells in Parkinson's disease). The absence of Lewy bodies in patients with parkin mutations suggests that parkin might be required for the formation of Lewy bodies**.** (Lang & Lozano, 1998)

.

PD is a destructive and complicated disease that interferes with movement more and more as time is passing by. Signs of the disease are vary divers, but they may consist of problems with swallowing, speech disorders, chewing, urinary problems or constipation, extreme sweating, skin problems, depression, other emotional changes, and difficulties with sleep or insomnia.

Nobody can predict which of these symptoms will have an effect on specific patient, and the strength of the symptoms is different from person to person. None of these secondary symptoms is mortal, although swallowing problems can lead to choking. The development of symptoms in Parkinson’s disease may take 20 years or more. In some people however, the disease advancement is greatly faster. In fact, it is one of the most usually used systems for explaining how the symptoms of PD progress are.

1.0.3 Treatments and solutions

Despite of there are currently two main types of treatment for PD: drug treatments and surgical treatments, but there are no treatments to curb the destruction of neurons in the substantia nigra. In prior studies of animal models of Parkinson's, researchers saw glimmers of success by injecting proteins that simulate neuronal growth into the animals' brains. A protein called glial-derived neurotrophic factor (GDNF) showed the most promise. In clinical trials, however, GDNF injections have failed to slow the course of the disease and in some cases; the injections have caused side effects such as severe weight loss. (Kim, et al., 2002)

In the new study of Gene therapy, monkeys were given a toxin (called MPTP) that destroys the same cells that are lost in Parkinson's disease. Several months after the monkeys developed Parkinson's-like symptoms, they received AAV-GDNF gene therapy by injections into the putamen, a brain region that connects to the subsantia nigra. (The substantia nigra is a brain structure located in the mesencephalon (midbrain) that plays an important role in reward, addiction, and movement).For more precise delivery, the researchers used a high pressure injection system, unlike previous studies where GDNF was injected under low pressure and allowed to spread by passive flow. As a control, other monkeys received saline injections which are gene therapy injections. Monkeys given the saline injections remained symptomatic, but monkeys that received the GDNF-virus injections got better. They showed an average 50 percent improvement on a symptom rating scale after 9 months, with smaller continued improvements out to 2 years. (Takagi, et al., 2005)

Japanese researchers have been capable to cure the signs of Parkinson's disease in monkeys by transplanting nerve cells isolated from embryonic stem cells into their brains, The finding is the world's first published success of its kind with a primate, pursuant to the

research team led by related professor Jun Takahashi of Kyoto University's Institute for Frontier Medical Sciences. After the transplant, those monkeys, which had been approximately incapable to move, demonstrated recovery in their symptoms to the position where they enable to walk on their own, rely on Embryonic stem cells, which have the potential to become almost any kind of tissue, are yielded from inside a blastocyst, which enhanced from a mammalian egg cell around a week after it is fertilized. Takahashi's scholars group used the embryonic stem cells to culture a cell crowd in which 35 percent of the cells were dopamine-generating neurons. Then these neurons were transplant into the four crab-eating monkeys, whose conditions were seen more than a one-year period. Based on the research, the monkeys illustrated reduced shaking of their organs half a year later. They had stayed virtually unmoved inside their cages all day long before the transplant, but the improvement of their symptoms finally makes them able to occasionally walk anent the cages. The research team confirmed that normal nerve cells had been produced in their brains. The detection could mark a great development for applying embryonic stem cells in clinical settings. (Takagi, et al., 2005)

Since Gene therapy is the use of a gene to change the function of cells or organs to improve or prevent disease. To transfer genes into cells, an inert virus is used to deliver the gene into a target cell. In this case, the glutamic acid decarboxylase (GAD) gene was used because GAD makes a chemical called GABA, a major inhibitory neurotransmitter in the brain that helps "quiet" excessive neuronal firing related to Parkinson's disease.Current pharmacological and surgical treatments for Parkinson's disease offer symptomatic improvements to those suffering from this incurable degenerative neurological disorder, but none of these has convincingly shown effects on disease progression. Novel approaches based on gene therapy have several potential advantages over conventional treatment modalities. These could be used to provide more consistent dopamine supplementation, potentially providing superior symptomatic relief with fewer side effects. More radically, gene therapy could be used to correct the imbalances in basal ganglia circuitry associated with the symptoms of Parkinson's disease, or to preserve or restore dopaminergic neurons lost during the disease process itself. (Kim, et al., 2002)

1.1 Problem Statement

Currently there are 91 structure of Amyloid beta A4 protein have been predicted in various species, but there is no presently concrete evidence that shows Amyloid beta A4 model in primates. However findings demonstrate that primates provide a close animal model for examining the early transcriptional and post transcriptional processing of APP (Amyloid precursor protein) that precedes it during aging in Alzheimer's diseases but in organisms like primates there isn't predicted model or sequence structure for amyloid β A4 protein.

1.2 Project Objectives

1. To develop a structural model for Amyloid beta A4 protein in primates.

2. To analyse the differences in the amino acid sequences and composition between different primates and human.

3. To compare 3D structure and investigation the implication of the differences in amino acid sequences of Amyloid beta A4 protein between human and primates.

1.3 Project Scope

The work will focus on Amyloid beta A4 protein information and analyzed using bioinformatics software: UniProt, PDB (protein data bank), BLAST, NCBI (National center for biotechnology information), and GenBank.

REFERENCES

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., et al. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallographica Section D: Biological Crystallography, 58(11), 1948-1954.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research, 25(17), 3389-3402.

Bard, F., Cannon, C., Barbour, R., Burke, R. L., Games, D., Grajeda, H., et al. (2000). Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease. Nature medicine, 6(8), 916-919.

Bowes, M. P., Masliah, E., Otero, D. A. C., Zivin, J. A., & Saitoh, T. (1994). Reduction of neurological damage by a peptide segment of the amyloid β/A4 protein precursor in a rabbit spinal cord ischemia model. Experimental neurology, 129(1), 112-119.

Cohen, F. E. & Sternberg, M. J. E. (1980). On the prediction of protein structure: the significance of the root-mean-square deviation. Journal of molecular biology, 138(2), 321-333.

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., & Barton, G. J. (1998). JPred: a consensus secondary structure prediction server. Bioinformatics, 14(10), 892-893.

Dev, C. O. G. (1974). 1. Kornberg RD, Thomas JO: Chromatin structure; oligomers of histones. Science, 184(2), 865-868.

Falicov, A., & Cohen, F. E. (1996). A surface of minimum area metric for the structural comparison of proteins. Journal of molecular biology, 258(5), 871-892.

Fraser, P. E., Nguyen, J. T., Inouye, H., Surewicz, W. K., Selkoe, D. J., Podlisny, M. B., et al. (1992). Fibril formation by primate, rodent, and Dutch-hemorrhagic analogs of Alzheimer amyloid. beta.-protein. Biochemistry, 31(44), 10716-10723.

Games, D., Adams, D., Alessandrini, R., Barbour, R., Borthelette, P., Blackwell, C., et al. (1995). Alzheimer-type neuropathology in transgenic mice overexpressing V717F β-amyloid precursor protein. 34(2), 324-453.

Gough, J., Karplus, K., Hughey, R., & Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. Journal of molecular biology, 313(4), 903-1154.

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis, 18(15), 2714-2723.

Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of chemical theory and computation, 4(3), 435-447.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. Journal of molecular biology, 233(1), 123-138.

Kim, J. H., Auerbach, J. M., Rodríguez-Gómez, J. A., Velasco, I., Gavin, D., Lumelsky, N., et al. (2002). Dopamine neurons derived from embryonic stem cells function in an animal model of Parkinson's disease. Nature, 418(6893), 50-56.

Lang, A. E., & Lozano, A. M. (1998). Parkinson's disease. New England Journal of Medicine, 339(15), 1044-1053.

Lee, S., Xue, Y., Hu, J., Wang, Y., Liu, X., Demeler, B., et al. (2011). The E2 domains of APP and APLP1 share a conserved mode of dimerization. Biochemistry, 50(24), 5453-5464.

Lustbader, J. W., Cirilli, M., Lin, C., Xu, H. W., Takuma, K., Wang, N., et al. (2004). ABAD Directly Links Aß to Mitochondrial Toxicity in Alzheimer's Disease. Science, 304(5669), 448-452.

Ramli, A. N. M., Mahadi, N. M., Rabu, A., Murad, A. M. A., Bakar, F. D. A., & Illias, R. M. (2011). Molecular cloning, expression and biochemical characterisation of a cold-adapted novel recombinant chitinase from Glaciozyma antarctica PI12. Microbial cell factories, 10(1), 94-134.

Ramli, A. N. M., Mahadi, N. M., Shamsir, M. S., Rabu, A., Joyce-Tan, K. H., Murad, A. M. A (2012). Structural prediction of a novel chitinase from the psychrophilic Glaciozyma antarctica PI12 and an analysis of its structural properties and function. Journal of computer-aided molecular design, 10(3), 1-15.

Sun, X., Zhulin, I. & Wartell, R. M. (2002). Predicted structure and phyletic distribution of the RNA-binding protein Hfq. Nucleic acids research, 30(17), 3662-3671.

Takagi, Y., Takahashi, J., Saiki, H., Morizane, A., Hayashi, T., Kishi, Y., et al. (2005). Dopaminergic neurons generated from monkey embryonic stem cells function in a Parkinson primate model. J Clin Invest, 115(1), 102-109.

Teller, J. K., Russo, C., Debusk, L. M., Angelini, G., Zaccheo, D., Dagna-Bricarelli, F., et al. (1996). Presence of soluble amyloid β–peptide precedes amyloid plaque formation in Down's syndrome. Nature medicine, 2(1), 93-95.

Wang, Y., & Ha, Y. (2004). The X-ray structure of an antiparallel dimer of the human amyloid precursor protein E2 domain. Molecular cell, 15(3), 343-353.