Role of Gene Expression in Neurodegenerative Disorders

Essay submitted for the partial fulfillment of Master Degree of Neuropsychiatry

submitted by

Mohamed Amr AbdelMonem Sherif
MB.B.Ch.
Ain Shams University

supervised by

Prof. Dr. Mohammad Ossama Abdulghani
Professor of Neuropsychiatry
Ain Shams University

Ass. Prof. Nagia Ali Fahmy
Assistant Professor of Neuropsychiatry
Ain Shams University

Dr. Lobna Mohammad El-Nabil
Lecturer of Neuropsychiatry
Ain Shams University

Faculty of Medicine
Ain Shams University
2007

دور التعبير الجينى فى الأمراض العصبية التنكسية

رسالة مقدمة توطئة للحصول على درجة الماجيستير فى فرع الأمراض العصبية و النفسية

مقدمة من

محمد عمرو عبد المنعم شريف
بكالوريوس الطب و الجراحة
جامعة عين شمس

تحت إشراف

أ.د. محمد أسامة عبد الغنى
أستاذ الأمراض العصبية و النفسية
كلية الطب
جامعة عين شمس

أ.م.د. ناجية على فهمى
أستاذ مساعد الأمراض العصبية و النفسية
كلية الطب
جامعة عين شمس

د. لبنى محمد النبيل
مدرس الأمراض العصبية و النفسية
كلية الطب
جامعة عين شمس

كلية الطب – جامعة عين شمس
2007

Gene expression is the process responsible for the different types of cells present in the human body, inspite that all of them contain the same genome. It is also responsible for cells responding in different ways to external stimuli. Gene expression is the process of transferring genetic information coded by the sequence of nucleotides within the DNA in the nucleus, to RNA that will cross the nuclear pore into the cytoplasm. The genetic information carried by that RNA, mRNA, will be translated into the sequence of amino acids within a polypeptide chain at the ribosomes. According to the sequence of the amino acids, the protein will fold into its three dimensional configuration that will determine its function. This may be followed by posttranslational modification. The final step is when the protein is no longer needed. It will be doomed to destruction and degradation, mostly through the ubiquitin-proteasome pathway *(Alberts et al, 2002).*

Neurodegenerative disorders represent an entity of neurological disorders that can present in different clinical pictures. However, they have a number of criteria in common. Firstly, a large group of that entity takes place because of formation of protein aggregates within neurons that ultimately leads to death of the neurons. Neurodegenerative disorders could be looked upon through the following framework; the

identity of the abnormally aggregating protein, the cause of its misfolding, causes of protein aggregation other than misfolding, the causes of failure of the ubiquitin-proteasome system to dispose the abnormally folded or aggregated protein, and the mechanism by which abnormally aggregated protein causes cellular damage. Accordingly, neurodegenerative disorders could be classified according to the abnormal protein aggregates, e.g. tauopathies, synucleinopathies and neuroserpinopathies (*Golbe, 2002*). Alzheimer's disease, as an example, is a tauopathy. It is diagnosed neuropathologically by the presence of amyloid plaques and neurofibrillary tangles (NFT) within the brain tissue. Amyloid plaques consist of beta-amyloid protein, while NFT consist of paired helical filaments (PHF) of hyperphosphorylated tau protein. It was found that the presence of mutations that result in increasing the level of beta-amyloid protein result in activation of certain kinases that lead to phosphorylation of tau proteins and thus NFT formation *(Avila, 2004).*

Our knowledge on the mechanisms of neurodegeneration is still in its infancy. For most neurodegenerative diseases, we can only guess as to why neurons ultimately die (*Bu et al, 2006*). This can take place due to simple clogging of the cell

by the aggregates or the aggregates rendering the cell more susceptible to oxidative stress. Another mechanism is through induction of apoptosis, which is enhanced by beta-amyloid in Alzheimer's disease and mutated huntingtin protein in Huntington's disease (*Gibson, 2001*).

Another criterion of neurodegenerative disorders is that they have long run-in period until sufficient protein accumulates, followed by a cascade of symptoms over 2-20 years, with increasing disability leading to death. Another character is that progressive degeneration occurs as a primary event long before symptoms develop and that it is selective, at least initially, for a particular neuronal pool (*Williams, 2002*).

The studying of different patterns of gene expression (i.e. transcriptome) gave us a major breakthrough in the understanding of the pathogenesis of neurodegenerative disorders including prion diseases. Determining the transcriptome represents a great challenge from a technical point of view. Firstly, the RNA will need to be extracted from the cells under study. Secondly, mRNA has to be analyzed in order to obtain useful information from all the available data that can throw light on new genes and proteins involved in the pathogenesis of the disease in question. This became possible

in the recent years through the technique of DNA microarrays (the so called Gene Chips). This tool is so powerful and the amount of data generated is huge that in order to make use of them, computer softwares has been designed and programmed for analysis and representation of the various data about gene expression, help to identify the function of the proteins encoded by the genes of interest through sequence comparison, 3-D structure analyses and comparison with other proteins of known function either within the same organism or among species *(Griffin, 2003*).

An important advance that was used to explore neurodegenerative disorders that I would like to highlight is animal models of diseases. The ability to develop models of different neurological disorders provided us by the opportunity to explore the gene families responsible for different diseases by identifying the abnormal gene(s) and protein(s) within these models. It also gave us the chance to try novel therapies. These models include mice (transgenic and knockout) and Drosophila (*Sang et al, 2005*).

Regulation of gene expression is essential in organ development and also in disease. It is this regulation that is responsible for cells of the embryo undergoing differentiation

to their final cell lines. Regulation can take place at many points on the pathway of information from DNA to proteins, the most common is at the level of transcription of DNA to mRNA. This is why knowing the expression pattern of different cells at variable times (i.e. the transcriptome), is essential as it tells us about the current proteins being in action in the cell at the mean time. This became possible in the recent years through the technique of DNA microarrays (*Alberts et al, 2002).*

The ability to interfere with or manipulate the pattern of gene expression has developed markedly in the past years with the development of techniques like RNA interference (RNAi) *(Wilton et al, 2005)* and exon skipping for manipulating alternative splicing. This helped markedly in our understanding of the molecular biology underlying neurodegenerative disorders. Also, potential drugs that has proven powerful in animal models are available that can modulate the rate of degradation of proteins *(Permanne et al, 2002).* Another class of drugs that is showing efficacy in animal models includes antisense techniques, the most powerful of which is RNAi. They play their role through decreasing the rate of synthesis of the proteins in question by

interfering with translation of mRNA to protein *(Dallas et al, 2006)*.

Aim of the work:

1- Highlighting the importance of gene expression and its regulation.

2- Gaining knowledge to the basics behind the tools to analyze and manipulate gene expression.

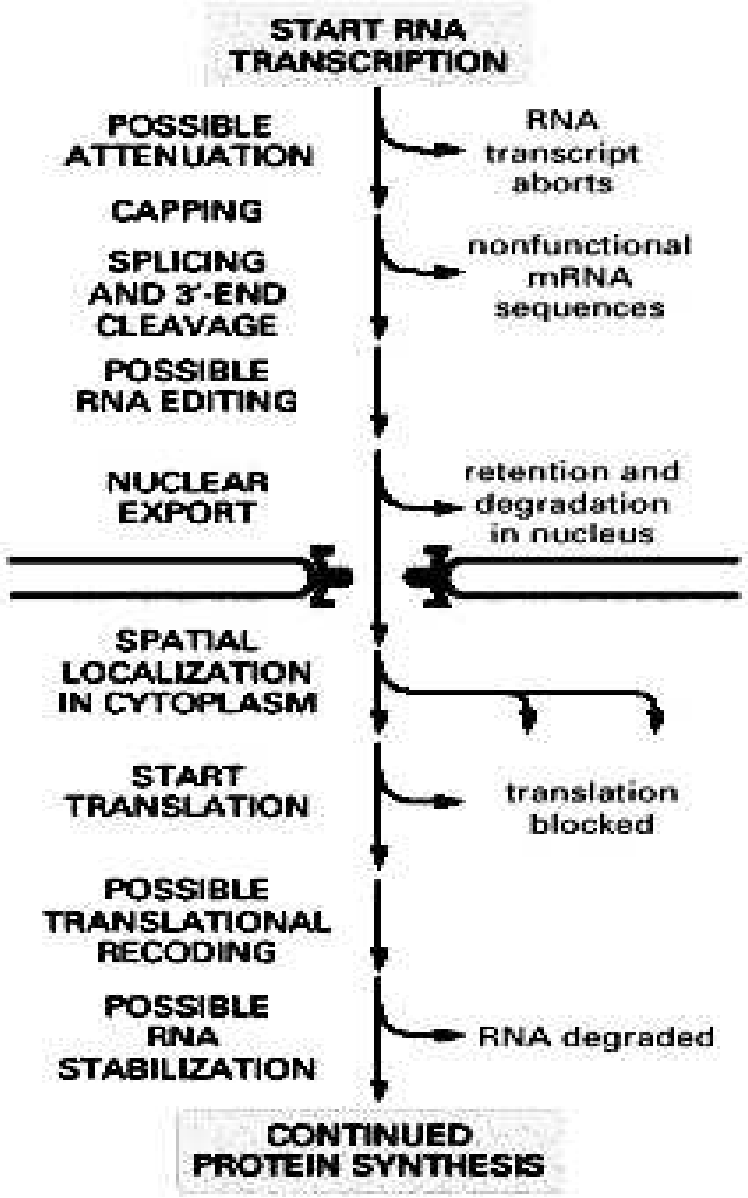
3- Exploring the pathogenesis of neurodegenerative disorders from a molecular point of view.

4- Identifying current and potential therapies for neurodegenerative disorders based on their molecular pathogenesis.

Molecular Biology of Gene Expression

The genome of a cell contains the information to make many thousands of different protein and RNA molecules in its DNA sequence. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control *(Alberts et al, 2002)*. Gene expression can be regulated at any of at least seven potential control steps; chromatin structure, initiation of transcription, processing of the transcript, transport of RNA to the cytoplasm, translation of mRNA, mRNA stability, protein activity stability (Fig. 1) *(Villard, 2004)*.

Fig. 1: Steps of gene expression and points of possible interference (Alberts et al, 2002).



Chromatin structure

Each chromosome consists of a single, enormously long linear DNA molecule associated with proteins that fold and pack the fine DNA thread into a more compact structure. The complex of DNA and protein is called chromatin. In addition to the proteins involved in packaging the DNA, chromosomes are also associated with many proteins required for the processes of gene expression, DNA replication, and DNA repair *(Rhodes, 1997)*.

Chromatin consists of nucleosomes connected by short lengths of DNA. The high-resolution structure of a nucleosome core particle revealed a disc-shaped histone core around which the DNA was tightly wrapped. In addition to its histone fold, each of the core histone subunits has a long N-terminal amino acid “tail”, which extends out from the DNA-histone core (Fig. 2). These histone tails are subject to several different types of covalent modifications, which control many aspects of chromatin structure *(Luger et al, 1997)*.

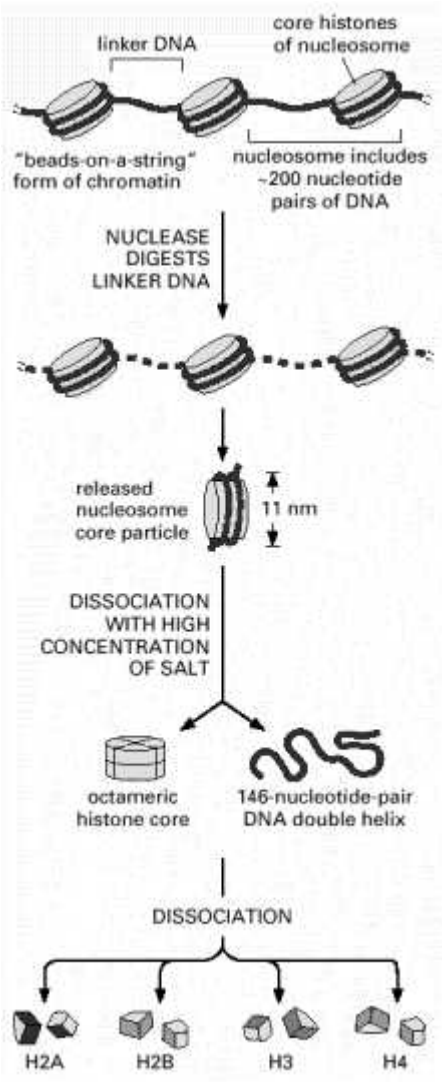


Fig. 2: Structural organization of the nucleosome (http://219.221.200.61/ywwy/zbsw(E)/edetail9)

The first step of gene expression regulation is implied by the fact that genes may exist in either of two structural conditions: a highly compact form that the packaged genes are transcriptionally silent (heterochromatin), and the "active" state in which the genes are much less compactly packged (euchromatin) so that they are transcribable. Histone

acetylation and the state of gene methylation are two mechanisms that are mainly implicated in the change of structure *(Emerson, 2002)*.

Initiation of transcription

As mentioned before, it is well demonstrated that the overwhelming majority of regulatory events occur at the initiation of gene transcription *(Mitchel et al, 1989)*.

Transcription (the transfer of genetic information from DNA to RNA) is initiated by the assembly of a transcription initiation complex composed of RNA polymerase II and other factors (general transcription factors). This multi-protein complex binds to a short DNA sequence called the minimal promoter, which often contains a conserved DNA sequence motif called the TATA box situated 20–30 nucleotides upstream of the transcription initiation site. The general transcription factors are characterized by their ability to control the activity of RNA polymerase II on the minimal promoter. These are TFIIB, D, E, F and H. They are called general as they share in the transcription of all genes *(Koleske et al, 1994)*.

For each gene or a set of genes, there exists a group of regulatory proteins (Regulatory transcription factors) that control the rate at which transcription takes place. Regulatory transcription factors (trans-acting elements) can positively or negatively affect the rate of transcription (activators (or enhancers) and repressors (or suppresors), respectively) by specific interaction with regulatory DNA sequences and by interactions with other proteins. Most of these transcription factors are activators. They would bind to specific sequences of DNA (cis-acting elements) that could be situated anywhere in relation to the gene. Then through the Mediator (which is another multisubunit protein) they would bind to and control the transcription initiation complex. Some of them even contain ligand binding sites, such as hormone binding sites that are essential for their activity *(Villard, 2004 and Kornberg, 2005)*.

The importance of a particular cis-acting element can vary greatly in different cell types and in response to different physiological stimuli because the transcription factors that bind to these elements may vary in abundance or in ability to function in different tissues and under different circumstances *(Levine et al, 2003)*.

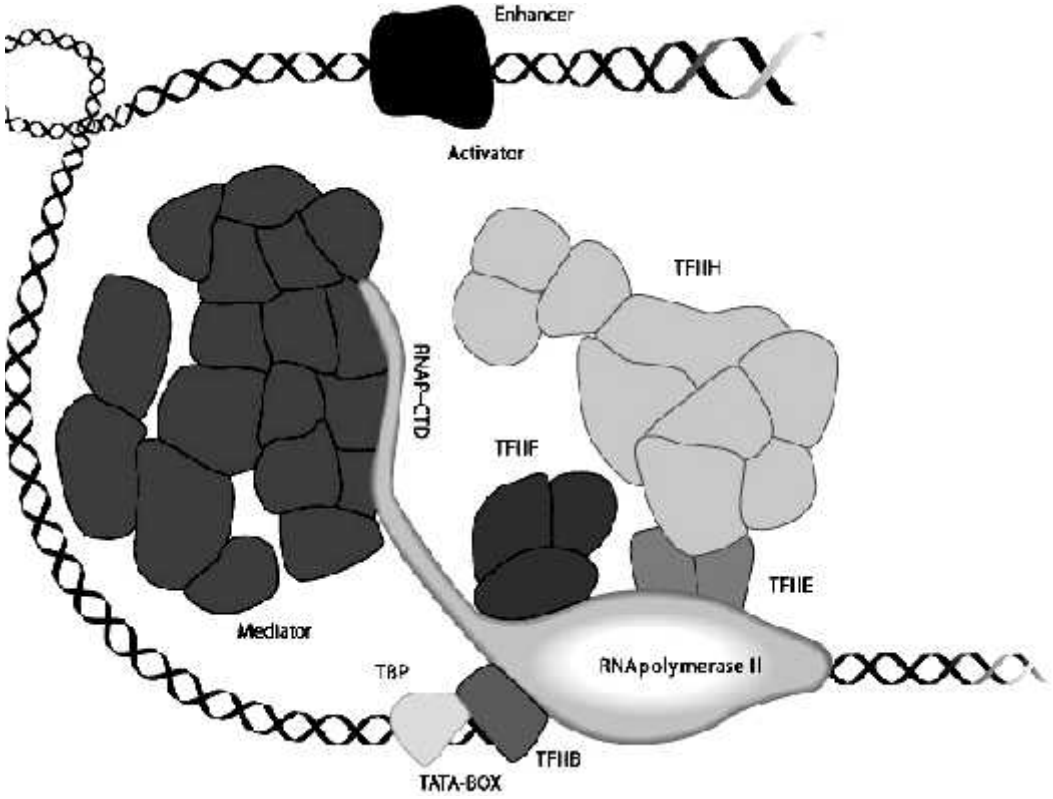


Fig. 3: A cartoon of a eukaryotic transcription initiation complex consisting of DNA, the general transcription factors TBP, TFIIB, E, F and H, Mediator, RNA polymerase II and a specific transcription factor binding to an enhancer element (http://nobelprize.org/nobel_prizes/chemistry/laureates/2006/chemadv06.pdf).

DNA looping is thought to allow gene regulatory proteins to bind at any of these positions to interact with the proteins that assemble at the promoter. The combination of regulatory proteins and their binding sites relative to the promoter are different for each gene (Fig. 3) *(Villard, 2004)*.