INTRODUCTION

The thalassemias are a diverse group of inherited hemolytic anemias characterized by reduced or absent production of one or more of the globin chains of hemoglobin ***(Weatherall and Clegg, 1981).*** Beta thalassemia is caused by any of more than 200 mutations that affect different levels of the beta-globin gene expression by a variety of mechanisms ***(Higgs et al., 2001).*** These mutations are ethnic-group specific; thus, only 4-6 account for the majority of beta-thalassemia alleles in each population ***(Brown et al., 1992).***

Beta-thalassemia is a major health problem in Egypt, with an estimated carrier state ranging from 1.3% ***(Sabry et al., 1981)*** to 10.2% ***(El-Beshlawy et al., 1999).*** More than 20 different mutations have been reported so far to cause beta-thalassemia in Egypt ***(Tadmouri and Gulen, 2003).*** Of these, 3 common Mediterranean mutations, namely IVS-I-110(G-A), IVS-I-6(T-C) and INS-I-1(G-A), accounted for the majority of beta thalassemia alleles in Egypt ***(Hussein et al., 1993; Rady et al., 1997; Waye et al., 1999; Omar et al., 2005; El-Gawhary et al., 2007).*** However, the relative frequencies of these common mutations varied among various studies. This may be due to the relatively small number of patients included in each study, or may reflect variation in geographic distribution of these mutations within the country. The spectrum of rarer mutations reported from different areas of Egypt was heterogeneous ***(Hussein et al., 2007).***

AIM OF THE WORK

The aim of this essay is:

* To review and summarize the results of molecular-based Egyptian studies of beta-thalassemia variants in order to obtain a better insight into the actual frequencies of these mutations in the country.

* To provide useful information for prospective studies aiming at cost-effective screening for beta-thalassemia mutations in Egypt as well as for genetic counseling, epidemiological and preventive measures.

Chapter I
HEMOGLOBIN STRUCTURE AND SYNTHSIS

A-Function of hemoglobin:

The main function of hemoglobin is to transport oxygen from the lungs to the tissues and CO_2 from the tissues back to the lungs. Hemoglobin also interacts with the physiologically important nitric oxide molecule (NO) to provide a mechanism for matching blood flow to metabolic demand in the peripheral tissues and for matching ventilation to perfusion in the lungs ***(Angelo et al., 2006)***. In addition, hemoglobin plays an important role in acid-base balance ***(Weatherall and Clegg, 2001a)***.

B-Structure of hemoglobin:

Hemoglobin is a conjugated protein which consists of specialized protein called globin that is tightly bound to 4 heme molecules ***(Ahmed et al., 2002)***.

Globin is a protein with 4 polypeptide chains joined together by non covalent bonds. Several different kinds of hemoglobin are normally found in humans. They vary in the primary structure of the peptide chains of globin. These include transient embryonic hemoglobins, Hb F, Hb A, Hb A_2 ***(Ahmed et al., 2002)***.

Table (1): Composition of hemoglobin in the human embryo, fetus and neonate.

Hemoglobin	Globin chains α-Globin gene cluster	β-Globin gene cluster	Gestation
Embryonic			
Hb Gower1	ζ_2	ε_2	From 3 weeks
Hb Gower2	α_2	ε_2	
Hb Portland	ζ_2	γ_2	From 4 weeks
Fetal			
HbF	α_2	γ_2	From 4 weeks
Adult			
Hb A	α_2	β_2	From 6-8 weeks
Hb A_2	α_2	δ_2	From 30 weeks

(Weatherall and Clegg, 2001b)

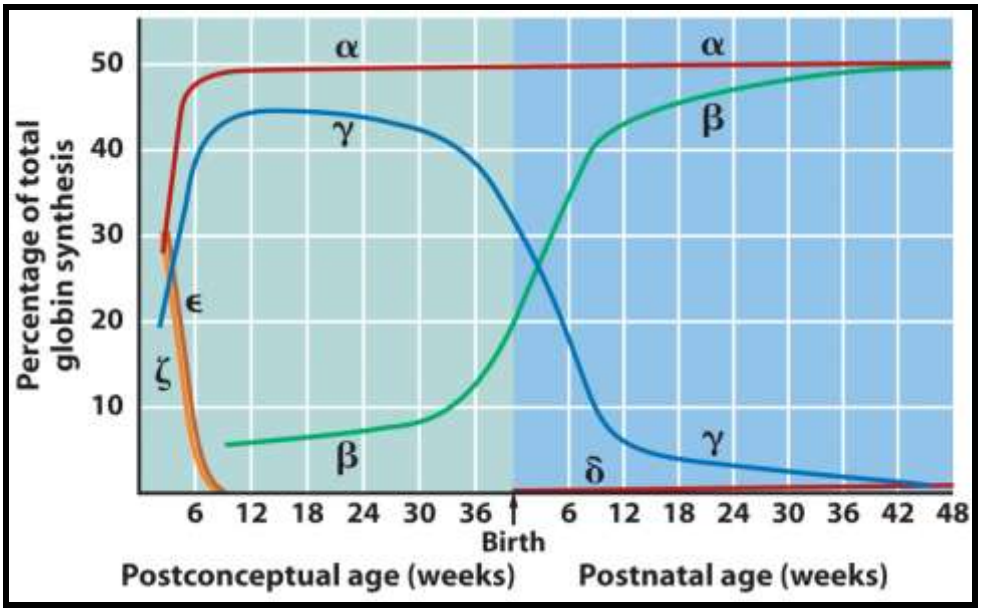

Figure (1): Human development of hemoglobin ***(Griffiths et al, 2008).***

Transient embryonic hemoglobins includes Hb Portland, Hb Gower 1 and Hb Gower 2 ***(Weatherall and Clegg, 2001b)***.

Fetal hemoglobin is the predominant hemoglobin of fetal life and comprises the major proportion of hemoglobin found at birth. It accounts for 1% of adult human hemoglobin ***(Manca and Masala, 2008)***.

Hemoglobin A is the major hemoglobin in adult (97%). Its globin comprises two α-chains (141 aminoacids) and two β-chains (146 aminoacids) ($\alpha_2\beta_2$). Each polypeptide chain has a helical structure and folded into 8 stretches labeled A to H creating a pocket inside it for heme binding ***(Manca and Masala, 2008)***.

The interior portion of the folded peptide chain is composed almost of non-polar aminoacids (hydrophobic). In contrast, polar aminoacids are located mostly on the surface of the molecule, where they can form hydrogen bonds with water (hydrophilic) ***(Manca and Masala, 2008)***.

The iron of each heme group is bound to the nitrogen atoms of histidine aminoacids, number 58 and 87 in α-chains and number 63 and 92 in β-chains.When hemoglobin is oxygenated, the bonds between iron and histidine molecules, number 87 in α-chain and number 92 in β-chain are displaced by oxygen ***(Yamada, 2004)***.

Hemoglobin A_2 accounts for about 2% of adult human hemoglobin ***(Nagel and Steinberg, 2001).***

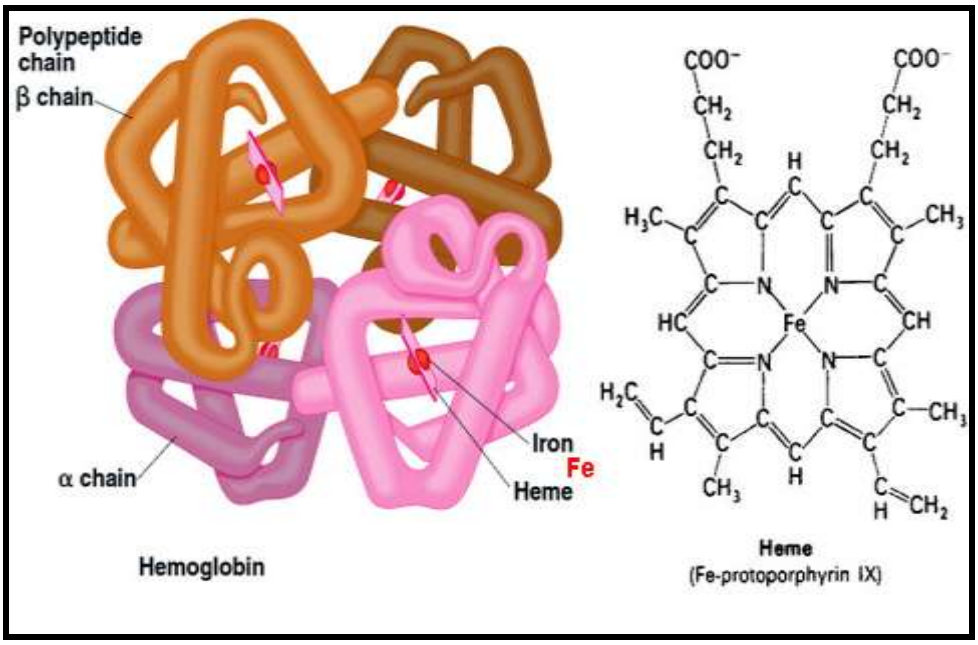

Figure (2): Structure of hemoglobin ***(Yamada, 2004).***

C-Catabolism of hemoglobin:

The average life span of the red blood cells is 120 days. At the end of that time, they are removed from circulation by the cells of reticuloendothelial (RE) system mostly present in liver, spleen and bone marrow, where they are hemolyzed and hemoglobin comes out. The globin molecule is hydrolyzed into free aminoacids while the heme ring is catabolized by heme oxygenase enzyme to biliverdin which is reduced to bilirubin by biliverdin reductase enzyme ***(Yamada, 2004)***.

Globin genes:

β-globin is encoded by a structural gene found in a cluster with the other β-like genes spanning 70 Kb on the short arm of chromosome 11.The cluster contains five functional genes, 5' ε -γ^G- γ^A –ψβ-δ-β 3', which are arranged in the order of their developmental expression ***(Thein, 2005)***.

Upstream of the entire β-globin complex is the locus control region (LCR), which is essential for the expression of all the genes in the complex. This region consists of five DNAase hypersensitive (HS) sites (designated HS1-5). The two extreme HS sites flanking the β-complex have been suggested to mark the boundaries of the β-globin gene domain ***(Thein, 2008)***.

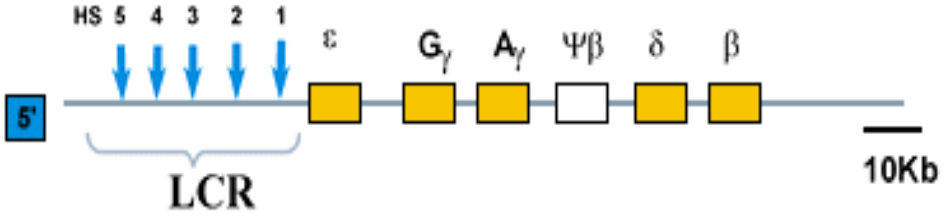

Figure (3): Organization of globin genes ***(Thein, 2005).***

The general structure of the β-globin gene is typical of the other globin loci. The transcribed region is contained in three exon separated by two introns or intervening sequence (IVS). The first exon encodes amino acids 1 to 29 together with the first two bases for codon 30, exon 2 encodes part of

residue 30 together with amino acids 31 to 104, and exon 3, amino acids 105 to 146 ***(Forget, 2001)***.

Conserved sequences important for gene function are found in the 5' promoter region, at the exon-intron junctions, and in the 3' untranslated region (3'UTR) at the end of mRNA sequences. The β-globin gene promoter includes 3 positive cis-acting elements: TATA box (position -28 to -31) a CCAAT box (position -72 to -76), and duplicated CACCC motifs (proximal at positions -86 to -90, and distal at position -101 to -105) ***(Marini, 2004)***.

These promoter elements are involved in the initiation of transcription and hence play an important role in the regulation of the structural genes ***(Marini, 2004)***.

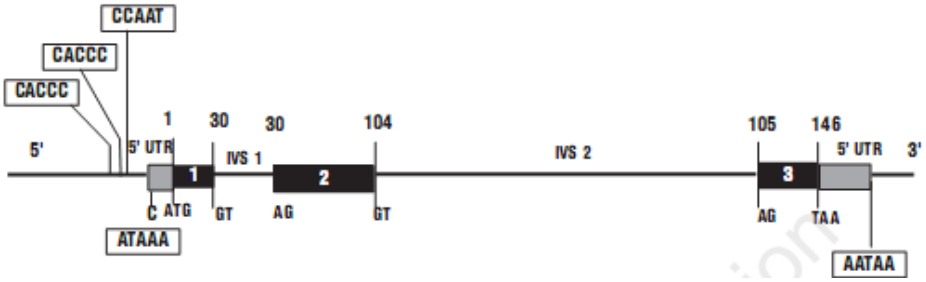

Figure (4): General structure of β-globin gene ***(Thein, 2005).***

Expression of globin genes:

Transcription and processing

When a globin gene is transcribed, messenger RNA (mRNA) is synthesized from one of its strand by the action of

RNA polymerase II. The primary transcript is a large mRNA precursor that contains both introns and exons, while in the nucleus, it undergoes a number of modifications first, the introns are removed and the exons are spliced together ***(Weatherall, 2001a)***.

The exon –intron junctions always have the sequence GT at their 5' end and AG at their 3' end, if there is a mutation at these sites, normal splicing cannot occur. The mRNA is modified at its 5' end by the addition of A cap structure, and at its 3' end by the addition of adenylic residues (poly A). The processed mRNA now moves into the cytoplasm to act as a template for globin chain production ***(Steinberg et al., 2001).***

Translation:

Amino acids are transported to the mRNA template on carriers called transfer RNAs, there are specific transfer RNAs for each aminoacid. The order of amino acids in a globin chain is determined by a triplet code i.e. three bases (codons) code for a particular amino acid ***(Forget et al., 2001).***

The transfer RNAs also contain three bases, anticodons, which are complementary to mRNA codons for particular amino acids. The transfer RNAs carry amino acids to the template, where they find the right position by codon-anticodon base pairing ***(Higgs et al., 2001a)***.

The mRNA is translated from the 5' end to the 3' end. The transfer RNAs are held in an appropriate steric conformation with the mRNA by the two subunits that make up the ribosomes ***(Steinberg et al., 2001).***

There are specific initiation (AUG) and termination (UAA, UAG, UGA) codons. When the ribosomes reach the termination codon, translation ceases, the completed globin chain is released, and the ribosomal subunits fall apart and are recycled ***(Steinberg et al., 2001).***

Individual globin chains combine with heme, which is synthesized through a separate pathway, and with themselves to form definitive hemoglobin ***(Weatherall, 2005b)***.

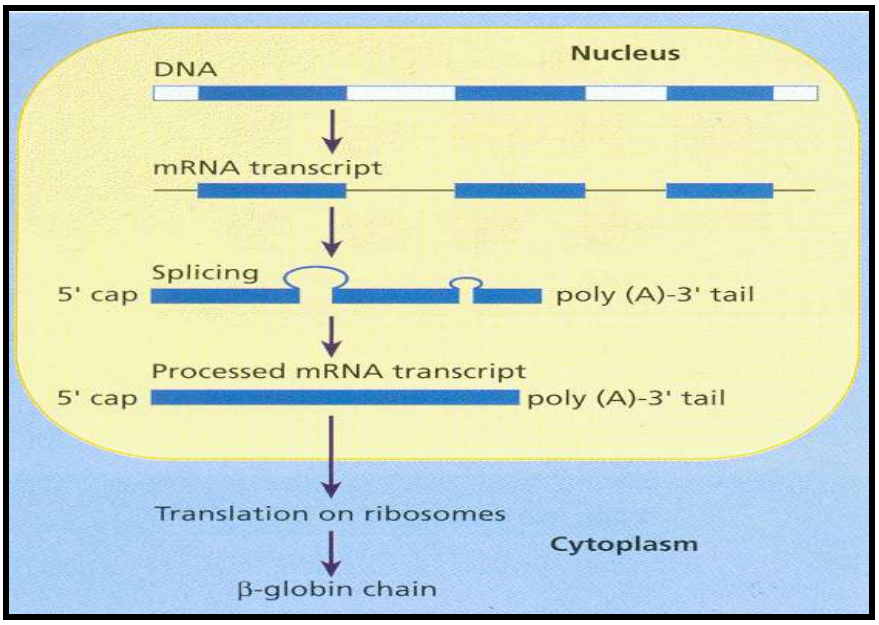

Figure (5) expression of globin genes ***(Rund and Rachmilewitz, 2005).***

Regulation of gene expression:

The genetic information is identical in any somatic cell, but the expression differs from one type of cells to the other (e.g. brain, liver). The type of response to the environment depends on the ability to alter and regulate the gene expression. This is accomplished through the interaction of certain proteins with promoters which are located at 5' adjacent to the coding gene segment ***(Weatherall, 2005a).***

Types of regulation:

I- Positive regulation: when a specific regulator element causes quantitative increase of gene expression.

II- Negative regulation: when a specific regulator element causes inhibition of gene expression.

III- Double negative regulation: when an effector inhibits the function of the negative regulator, hence releasing the positive regulatory effect ***(Lewis et al., 2000)***.

Types of gene response:

I- Type A response: depends on the continued presence of the inducing codon; when present we get gene expression, when removed, gene expression stops and when it reappears, the increase of gene expression returns ***(Lewis et al., 2000)***.

II- Type B response: when there is increased gene expression but it is transient and stops even with the continuation of the inducing codon and the cell recovers ***(Lewis et al., 2000)***.

Chapter II

THE THALASSEMIAS

The thalassemias are a heterogeneous group of inherited disorders of globin synthesis, all of which result from a reduced rate of production of one or more of the globin chains of hemoglobin. The result is imbalanced globin chain production, ineffective erythropoiesis, hemolysis and variable degree of anemia ***(Omar et al., 2005)***.

They are divided into α-and β-thalassemias according to which globin chain is produced in a reduced amount ***(Weatherall and Clegg, 2001b)***.

The main genetics variants of β-thalassemia are:

I- β-thalassemia which is divided into two main varieties: β^0-thalassemia, where there is no β-chain production and β^+-thalassemia, where there is a partial deficiency of β-chain production ***(Steinberg et al., 2001)***.

II- β-thalassemia associated with β-chain structural variants.

III- Variants of β-thalassemia.

IV- δβ-thalassemia.

V- εγδβ-thalassemia.

VI- Hereditary persistence of fetal hemoglobin (HPFH)

(Olivieri, 1999)

I-β-thalassemia:

Geographic distribution and epidemiology:

Thalassemia is the most common genetic disorder allover the world. Around 3% of the world population carries genes for β-thalassemia ***(Omar et al., 2005)***.

β-thalassemia is distributed widely in Mediterranean populations, Middle East, parts of India and Pakistan, and throught Southeast Asia. The disease is common in parts of the southern Soviet republics and in the People's Republic of China. β-thalassemia is rare in Africa, except for isolated pockets in West Africa, notably Liberia, and in parts of North Africa ***(Weatherall and Clegg, 2001a)***.

The most common forms of thalassemia are those that are prevalent in the malarial tropical and sub-tropical regions where a few mutations have reached high gene frequencies because of the protection they provide against malaria. The epidemiology of the disease; however is changing due to a fall in total birth rate, prevention programs and recent population movements ***(Flint et al., 1998)***.

Figure (5): Geographic distribution of β-thalassemia ***(Weatherall and Clegg, 2001b)***.

Molecular pathology of β-thalassemia:

β-thalassemia is extremely heterogeneous at the molecular level. More than 200 different mutations have been found in association with the β-thalassemia phenotype. They may involve any step in globin chain production: transcription, processing, translation or post-translational stability of the β-globin gene product ***(Weatherall, 2001b)***.

A-Transcription:

Point mutations within promoter sequence tend to reduce binding of RNA polymerase, so reducing the rate of mRNA transcription to 20-30% of normal. The result is a moderate decrease of β-globin chain output and hence, in a mild phenotype (β^+- thalassemia) ***(Weatherall and Clegg, 2001a)***.

B- Processing

A wide variety of mutations interfere with processing of the primary mRNA transcript. Those within introns or exons, or at their junctions, interfere with the mechanism of splicing the exons together after the introns have been removed ***(Weatherall and Clegg, 2001a)***.

Single base substitutions at the invariant GT or AG sequences at intron-exon junctions prevent splicing altogether and cause β^0- thalassemia ***(Weatherall and Clegg, 2001a)***.

Other mutations involve consensus sequences adjacent to the introns leading to formation of alternative splicing sites which compete with normal splicing sites so that normal and abnormal mRNA species are synthesized ***(Weatherall and Clegg, 2001a)***.