

# MOLECULAR GENETICS OF A-THALASSEMIA

*Titus H.J. Huisman*

Remarkable progress has been made in our understanding of the molecular biology of the red cell during the past 10–15 years; among others, the locations and sequences of the various globin genes were established. Two  $\alpha$ -globin genes ( $\alpha_2$  and  $\alpha_1$ ), one embryonic globin gene ( $\zeta$ ), at least three pseudo genes, and one gene of undetermined function ( $\psi_1$ ) are located at the tip of chromosome #16 (order: Telomere- $\zeta$ - $\psi_1$ - $\alpha_2$ - $\alpha_1$ -q1-centomere), while the  $\beta$ -globin genes are found on the short arm of chromosome #11 (order: 5'-e-G<sub>g</sub>-A<sub>g</sub>-y<sub>b</sub>-d-b-3'). The coding regions (exons 1, 2, and 3) of the  $\alpha_2$ - and  $\alpha_1$ -globin genes have identical sequences, and a mutation in either one of these genes results in the synthesis of an  $\alpha^X$  chain. Presently, nearly 200  $\alpha$ -chain variants with only one amino acid substitution have been described. Studies by Liebhaber et al<sup>(1,2)</sup> showed that the relative quantities of  $\alpha_2$ -mRNA in reticulocytes was 2.5- to 3-times higher than that of  $\alpha_1$ -mRNA, which led to the conclusion that the  $\alpha_2$  gene is the major gene and the  $\alpha_1$  the minor gene. Molchanova et al<sup>(3)</sup> reevaluated this through an analysis of the relative quantities of the  $\alpha_2$ - and  $\alpha_1$ -mRNAs in more than 60 patients with a chain variants. They developed a PCR-based method for mRNA quantitation. Amplification of the corresponding cDNAs is based on sequence differences in the 3'UTR (untranslated region); selection of different reverse primers and a common forward primer allowed the formation of amplification products that were 298 bp for  $\alpha_1$  and 295 bp for  $\alpha_2$ . Digestion with the restriction enzyme *Sly I* facilitated the separation of the PCR products because a restriction site is present in the  $\alpha_1$  fragment (giving a shorter 283 bp fragment) and not in the  $\alpha_2$  fragment. Separation was obtained on a 6% polyacrylamide-urea gel, and scanning of the intensities of the bands provided the relative quantities of the  $\alpha_2$ - and  $\alpha_1$ -mRNAs. Furthermore, HPLC methodology was used to quantify  $\alpha$  chain variants in red cell lysates.<sup>(4)</sup> The data confirm an mRNA ratio of about 2.7 ( $\alpha_2/\alpha_1$ ) for normal persons and for subjects with an  $\alpha$  chain variant. However, the percentages of stable  $\alpha_2$  abnormal Hbs (n = 9) averaged 23.5±1.9% (range 18–28%) and those for stable  $\alpha_1$  abnormal Hbs (n = 7) averaged 19.7±1.6% (range 16.9–22.5%). These percentages correspond to a ratio of  $\alpha_2$  to  $\alpha_1$  of 1.19:1 at the protein level. Thus, the high relative  $\alpha_2$ -mRNA level is about twice the relative level of the  $\alpha_2$  protein, which suggests either a less efficient translation of the  $\alpha_2$ -mRNA or a decreased stability.

$\alpha$ -Thalassemia ( $\alpha$ -thal) is caused by underproduction of  $\alpha$  chains and has the highest frequency of all genetic disorders. It is observed mainly in Mediterranean populations, the Middle East, India, the various East Asian countries, Oceania, and also in the Blacks of Africa and the Americas. Nearly 50 different abnormalities are known to exist. The genetics of  $\alpha$ -thal is complex because of the duplication of the  $\alpha$ -globin gene. An adult with four  $\alpha$ -globin genes ( $\alpha\alpha/\alpha\alpha$ ) has a balanced chain synthesis. Deletion ( $-\alpha/\alpha\alpha$ ) or inactivation [ $\alpha(T)\alpha/\alpha\alpha$ ] of one  $\alpha$ -globin gene results in an  $\alpha$ -thal-2 heterozygosity characterized by a slight excess of non- $\alpha$  chain, and a mild microcytosis or hypochromia. These abnormalities become more pronounced when additional  $\alpha$ -globin genes are deleted or deactivated; subjects with the  $\alpha$ -thal-2 homozygosity ( $-\alpha/-\alpha$ ) and  $\alpha$ -thal-1 heterozygosity ( $--/\alpha\alpha$ ) have a mild anemia, distinct microcytosis and hypochromia. In the

newborn, the excess  $\gamma$  chain forms Hb Bart's or  $\gamma_4$  (about 5% in babies with  $-a/-a$  or  $--/aa$ ), while the comparable adult form (Hb H or  $\beta_4$ ) is present in only minute quantities in adults with these deficiencies. The loss of three  $\alpha$ -globin genes ( $--/a$ ), either by deletion or deactivation, results in Hb H disease, a distinct microcytic hypochromic anemia (Hb 8-10 g/dl) with Hb Bart's at birth of 15–20% and with Hb H in the adult of 5–25%. Finally, homozygosity for an  $\alpha$ -thal-1 determinant is known as hydrops fetalis ( $--/--$ ), which is not compatible with life.

$\alpha$ -Thal is most often caused by deletion of one or two  $\alpha$ -globin genes. There are at least eight different types of  $\alpha$ -thal-2 ( $-a/$ ), of which the 3.7 and 4.2 kb deletions are the most common. Crossovers between misaligned chromosomes result in the  $-a(-3.7$  kb) allele (crossover between homologous subfragments, the Z boxes) or the  $-a(-4.2$  kb) allele (crossover between the comparable X boxes) and in chromosomes with three  $\alpha$  genes [the so-called  $aaa(\text{anti } 3.7)$  and  $aaa(\text{anti } 4.2)$  alleles]. The number of larger deletions involving both  $\alpha$ -globin genes ( $\alpha$ -thal-1) is at least 16. Furthermore, five deletions that include the regulatory segment HS-40 but not the  $\alpha_2$ - and  $\alpha_1$ -globin genes result in similar  $\alpha$ -thal-1 types. Detection of these deletions is by hematological analysis and by gene mapping. Recently, PCR-based methodology has been developed to characterize the 3.7 and 4.2 kb deletions<sup>(5)</sup> and a few  $\alpha$ -thal-1 deletions.<sup>(6,7)</sup> In addition, there are nondeletional mutants that may affect RNA processing, RNA translation, or cause the synthesis of an unstable Hb variant which is destroyed by proteolysis and is not incorporated in a Hb tetramer. As many as 20 different nondeletional  $\alpha$ -thal mutations have been detected, mostly affecting the  $\alpha_2$ -globin gene.

Clinically the most important is Hb H disease, resembling thalassemia intermedia, although there is a significant variability in severity. The condition is marked by anemia, jaundice, hepatosplenomegaly, Hb H ( $\beta_4$ ) in peripheral blood, infections, leg ulcers, gallstones, and acute hemolysis after treatment with certain drugs (for further details see Refs. 8, 9, and 10). Hb H disease is most often the result of combinations of an  $\alpha$ -thal-1 [ $--/$ , mainly the SEA type in the Far East; the MED-I and  $-(a)20.5$  kb deletions in the Mediterranean area] with an  $\alpha$ -thal-2 ( $-3.7$  or  $-4.2$  kb). The interaction of  $\alpha$ -thal-1 ( $--/$ ; the SEA type) with a nondeletional form [Hb Constant Spring (Hb CS) caused by a mutation in the terminating codon,  $\alpha_{142}(\underline{TAA}\underline{ECAA})$ ] is commonly found in Chinese patients. Homozygosity for a nondeletional  $\alpha$ -thal-2 such as the poly A mutation  $AATAAA\overline{EAATAAG}$  (the PA-1 type) or  $\overline{EAATGAA}$  (the PA-2 type) and the 5 nucleotide (nt) deletion at the IVS-I donor splice site ( $GAGGTGAGG\overline{EGAGG}$ ----) can also cause Hb H disease. These conditions and combinations of deletional  $\alpha$ -thal ( $--/$ ) with one of the nondeletional  $\alpha$ -thal-2 determinants [ $--/a(\text{PA-1})a$  or  $--/a(\text{PA-2})a$ ;  $--/a(-5\text{nt})a$ ] are mainly found in the population of the Eastern Mediterranean and the Middle East.

Recently developed PCR-based methodology has proven to be rather helpful in the study of the various forms of Hb H disease. The (functional) loss of one or more  $\alpha$ -globin genes by deletion or mutation may lead to a change in the ratio between the  $\alpha_2$ - and  $\alpha_1$ -mRNAs and/or in that between the  $\beta$ - and  $(\alpha_2+\alpha_1)$ -mRNAs. Both ratios can be determined from a mixture of PCR products being synthesized in the same tube because the  $\alpha_2$ ,  $\alpha_1$ , and  $\beta$  primers can be added simultaneously. Table 1 summarizes some results. The differences in the two are quite diagnostic. The  $\alpha$ -thal-2 conditions ( $-a/aa$  and  $-a/-a$ ) are easily recognized when the two ratios are compared with those for normal persons.

Both ratios are also lower in persons with a heterozygosity for Hb CS, which appears promising for the differential diagnosis of Hb H disease. The normal  $\alpha_2/\alpha_1$  ratio in  $\alpha$ -thal-1 heterozygotes ( $--/aa$ ) is expected because mRNA is produced only by the normal chromosome, while the  $a/b$  ratio is lowered by nearly 50%.

Table 1. The  $\alpha_2/\alpha_1$ - and  $\alpha/\beta$ -mRNA ratios in some  $\alpha$ -thalassemias (average values only).

Condition	$\alpha_2/\alpha_1$ Ratio	$\alpha/\beta$ Ratio
Normal ( $\alpha\alpha/\alpha\alpha$ )	2.5 - 2.7	4.2 - 4.6
Triplication ( $\alpha\alpha/\alpha\alpha\alpha$ )	4.0	5.0
$\alpha$ -Thal-2 trait ( $-\alpha/\alpha\alpha$ )	1.5 - 1.6	3.8 - 4.0
$\alpha$ -Thal-2 homozygote ( $-\alpha/-\alpha$ )	No $\alpha^2$	1.2 - 1.4
$\alpha$ -Thal-2 trait (Hb CS or $aCSa/aa$ )	1.9 - 2.1	2.9 - 3.0
$\alpha$ -Thal-1 trait ( $--/\alpha\alpha$ )	2.5 - 2.7	2.4 - 2.6
Hb H disease ( $--/-\alpha$ )	0	0.9 - 1.1
Hb H disease ( $--/\alpha^{CS}\alpha$ )	0.15- 0.25	0.5 - 0.6

Application of this technique to RNA isolated from reticulocytes of patients with Hb H disease with different genotypes provided data that are summarized in Figure 1 (from Ref. 11).

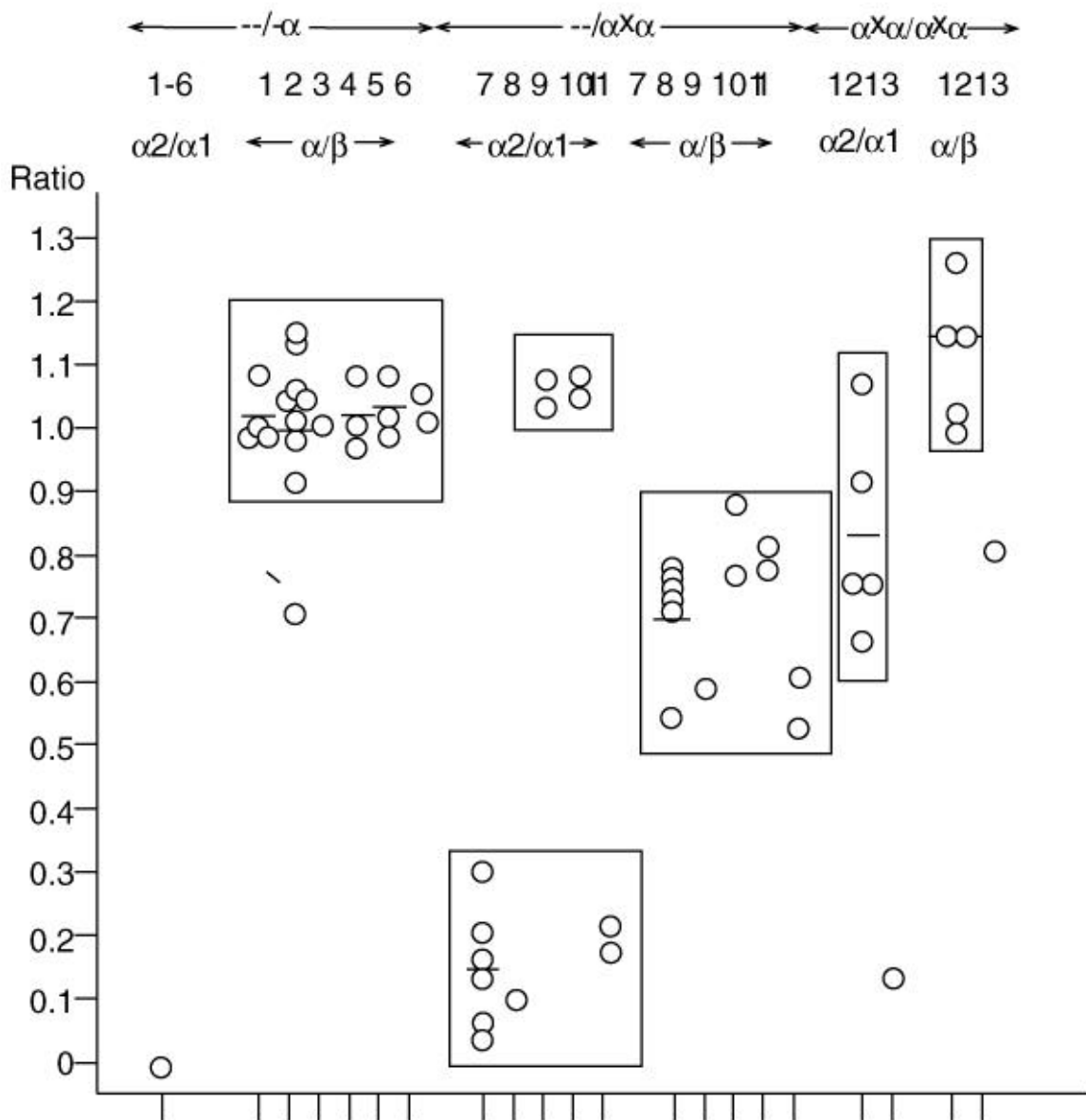


Figure 1. 1 through 6: Hb H disease due to  $\alpha$ -thal-1 combined with deletional  $\alpha$ -thal-2 ( $--/\alpha$ ). 1:  $--(\text{MED-II})/\alpha(3.7)$ ; 2:  $--(\text{MED-I})/\alpha(3.7)$ ; 3:  $--(\text{MED-II})/\alpha(4.2)$ ; 4:  $-(\alpha)20.5/\alpha(3.7)$ ; 5:  $--(\text{SEA})/\alpha(3.7)$ ; 6:  $--(\text{SEA})/\alpha(4.2)$ .  
 7 through 11: Hb H disease due to  $\alpha$ -thal-1 combined with nondeletional  $\alpha$ -thal-2 [ $--/\alpha(\text{T})\alpha$ ]. 7:  $--(\text{MED-I})/\alpha(-5\text{nt})\alpha$ ; 8:  $-(\alpha)20.5/\alpha(-5\text{nt})\alpha$ ; 9:  $--(\text{MED-I})/\alpha(\text{PA})\alpha$ ; 10:  $-(\alpha)20.5/\alpha(\text{PA})\alpha$ ; 11:  $--(\text{SEA})/\alpha(\text{CS})\alpha$ .  
 12 and 13: Hb H disease due to homozygosity for a nondeletional  $\alpha$ -thal-2 [ $\alpha(\text{T})\alpha/\alpha(\text{T})\alpha$ ].  
 12:  $\alpha(\text{PA})\alpha/\alpha(\text{PA})\alpha$ ;  
 13:  $\alpha(-5\text{nt})\alpha/\alpha(-5\text{nt})\alpha$ .

Three major types of Hb H disease can be recognized. Patients with deletional types ( $--/\alpha$ ), specified in the legend of the figure, have no  $\alpha 2$ -mRNA and an  $\alpha/b$  ratio of one (irrespective of the types of  $\alpha 1$ - or  $\alpha 2$ -thal). Translation of the mRNA from the one

remaining  $\alpha$ -globin gene must be most efficient because the patients are able to maintain a Hb level of 8–10 g/dl. The  $\alpha/b$ -mRNA ratio for the nondeletional Hb H patients [--/ $\alpha(T)$ ], five different genotypes, averages ~0.7. Furthermore, the  $\alpha_2/\alpha_1$ -mRNA ratio is above zero at 0.05–0.3 (T = the 5 nt deletion or the termination codon mutation) and at 1.0–1.1 (T = a poly A mutation). The data for the  $\alpha(T)\alpha$  homozygotes are comparable: Low  $\alpha_2/\alpha_1$ - and  $\alpha/b$ -mRNA ratios are observed if T is the 5 nt deletion at the IVS-I donor splice site, and values between 0.7 and 1.1 for  $\alpha_2/\alpha_1$ -mRNA and 0.9–1.3 for  $\alpha/b$ -mRNA in the patients who are homozygous for a poly A mutation. These data are of considerable diagnostic significance and will aid in the search for nondeletional  $\alpha$ -thal.

The low  $\alpha/b$ -mRNA ratio in Hb H patients with the genotype --/ $\alpha(T)\alpha$  suggests a more severe disease; indeed, a more significant anemia is observed in these patients with a frequent need for blood transfusions and with higher Hb H values. Surprising are the differences in the  $\alpha_2/\alpha_1$  ratios for Hb H patients with the nondeletional  $\alpha$ -thal-2 alleles. The presence of a small amount of  $\alpha_2$ -mRNA in patients with the genotype --/ $\alpha(-5nt)\alpha$  or  $\alpha(-5nt)\alpha/\alpha(-5nt)\alpha$  indicates that the TGAGG deletion at the IVS-I donor splice site (AG\_GTGAGGCTCCCÆAG\_GCTCCC) decreases but does not completely eliminate the splicing. The high  $\alpha_2/\alpha_1$ -mRNA ratio of about one for the Hb H patients with the poly A mutation [--/ $\alpha(PA-2)\alpha$ ] or with a homozygosity for this mutation has been evaluated with additional RT-PCR experiments, using reverse primers for  $\alpha_2$  and  $\alpha_1$  that are located 64 bp downstream of the major polyadenylation signal.<sup>(12)</sup> The resulting fragment of 411 bp was found for the  $\alpha_2$ -mRNA but the corresponding 420 bp fragment for the  $\alpha_1$ -mRNA was not observed. This unusual, elongated mRNA is present in the reticulocytes of subjects carrying a polyadenylation signal mutation, while minute quantities were seen in normal adults and larger quantities in persons with a terminating codon mutation. Indeed, in patients with --/ $\alpha(PA-1)\alpha$  or  $\alpha(PA-1)\alpha/\alpha(PA-1)\alpha$ , the  $\alpha_2$ -mRNA is of the elongated type only; this mRNA likely uses the cryptic poly A sequence at 1048 bp past the terminating codon as its poly A site. In patients with a terminating codon mutation, this elongated mRNA is slightly different; it uses the TAA sequence of the normal poly A site as terminating codon and the cryptic poly A sequence at 1048 bp as the polyadenylation site.

The rapid advance in the mRNA study is a remarkable development; the RT-PCR methodology makes it possible to analyze even the smallest quantity of RNA that can be isolated from the few reticulocytes of a small blood sample from a normal person. This new technique provides new insight into the molecular pathology of the  $\alpha$ -thalassemias and has potential for a rapid and more detailed diagnosis. However, the data obtained through mRNA analyses will give at present preliminary diagnoses only which should be confirmed with DNA methodology.

## References

1. Liebhaber SA, Kan YW: Differentiation of the mRNA transcripts originating from the  $\alpha_1$ - and  $\alpha_2$ -globin loci in normals and  $\alpha$ -thalassemics. *J Clin Invest* 68:439, 1981.
2. Liebhaber SA, Cash FE, Ballas SK: Human  $\alpha$  globin gene expression. The dominant role of the  $\alpha_2$ -locus in mRNA and protein synthesis. *J Biol Chem* 261:15327, 1986.

3. Molchanova TP, Pobedimskaya DD, Huisman THJ: The differences in quantities of  $\alpha 2$ - and  $\alpha 1$ -globin gene variants in heterozygotes. *Br J Haematol* 88:300, 1994.
4. Huisman THJ (ed): *The Hemoglobinopathies*, Methods in Hematology, Vol 15, Churchill Livingstone, Edinburgh, 1986.
5. Baysal E, Huisman THJ: Detection of common deletional  $\alpha$ -thalassaemia-2 determinants by PCR. *Am J Hematol* 46:208, 1994.
6. Bowden DK, Vickers MA, Higgs DR: A PCR-based strategy to detect the common severe determinants of  $\alpha$  thalassaemia. *Br J Haematol* 81:104, 1992.
7. Dode C, Rochette J, Krishnamoorthy R: Locus assignment of human  $\alpha$ -globin mutations by selective amplification and direct sequencing. *Br J Haematol* 76:275, 1990.
8. Weatherall DJ, Clegg JB: *The Thalassaemia Syndromes*, 3rd edition, Blackwell Scientific Publications, Oxford, 1981.
9. Bunn HF, Forget BG: *Hemoglobin: Molecular, Genetic and Clinical Aspects*, WB Saunders Company, Philadelphia, 1986.
10. Higgs DR, Weatherall DJ (eds): *The Haemoglobinopathies*, Baillière's Clinical Haematology, Vol 6, WB Saunders Company, London, 1993.
11. Smetanina NS, Öner C, Baysal E, Öner R, Bozkurt G, Altay Ç, Gürgey A, Adekile AD, Gu L-H, Huisman THJ: The relative levels of  $\alpha 2$ -,  $\alpha 1$ -, and  $\alpha$ -mRNA in Hb H patients with different deletional and nondeletional  $\alpha$ -thalassaemia determinants. *Biochim Biophys Acta*, in press.
12. Molchanova TP, Smetanina NS, Huisman THJ: A second, elongated,  $\alpha 2$ -globin mRNA is present in reticulocytes from normal persons and subjects with terminating codon or poly A mutations. *Biochem Biophys Res Commun* 214:1184, 1995.