alpha-Thalassemia Caused by a Large (62 kb) Deletion Upstream of the Human 
alpha Globin Gene Cluster

V.J. Buckle, and D.R. Higgs

We describe a family in which alpha-thalassemia occurs in association with a deletion of 62 kilobases from a region upstream of the alpha globin genes. DNA sequence analysis has shown that the transcription units of both alpha genes downstream of this deletion are normal. Nevertheless, they fail to direct alpha globin synthesis in an interspecific hybrid containing the abnormal (alpha)RA chromosome. It seems probable that previously unidentified positive regulatory sequences analogous to those detected in a corresponding position of the human beta globin cluster are removed by this deletion.

THE HUMAN alpha globin gene cluster lies near the tip of the short arm of chromosome 16 within band p13.3.1,2 The complex includes two alpha genes (alpha-1 and alpha-2), an embryonic alpha-like gene (alpha-2), several pseudogenes (alpha-1, alpha-2, and alpha-3) and a gene of undetermined function (theta-1) in order (5'-alpha-2-theta-1-alpha-2-alpha-1-theta-3').3,7 Alpha-thalassemia most frequently results from deletions including either one (denoted alpha) or both (denoted alpha') alpha genes from one chromosome, although a variety of nondeletion mutations (denoted alphaR or alpha') have also been described. In all of these determinants of alpha-thalassemia, the reduced alpha globin chain synthesis can be explained in terms of our current understanding of globin gene expression.8

We have recently identified an English individual (R.A.) with alpha-thalassemia in which the molecular basis cannot be so readily explained. DNA analysis has demonstrated a large (62-kilobase [kb]) deletion affecting one chromosome denoted (alpha)RA, spanning from coordinate +10 between the theta-2 and theta-1 genes to a position 52 kb upstream of the theta-2 messenger RNA (mRNA) CAP site (see Fig 1). However, both alpha genes downstream of the deletion in the chromosome remain intact.9 Therefore, it is not clear whether the deletion is primarily responsible for the alpha-thalassemia phenotype or if it is merely a coincidental mutation either in cis or trans to another downregulating mutation.

To establish the relationship of this novel deletion to the associated alpha-thalassemia we have determined the phenotype of other family members with the alpha/alphaR genotype. Furthermore, the structurally normal (alpha) and abnormal [(alpha)RA] chromosomes 16 derived from the propositus were each isolated in mouse x human somatic cell hybrids for detailed structural and functional studies. Finally, the alpha globin genes located downstream of the deletion in the abnormal (alpha)RA chromosome were cloned and sequenced to search for mutations that might explain the phenotype, but none was found. Together, these studies strongly suggest that the 62-kb deletion is primarily responsible for downregulating the intact alpha genes on the (alpha)RA chromosome.

MATERIALS AND METHODS

Hematologic analysis. Full blood counts, hemoglobin electrophoresis, and HbH preparations were performed using standard methods.10,11 and alpha/beta globin chain synthesis ratios were measured as previously described.12 DNA analysis. Blot hybridization studies were performed using standard methods.13,14 Previously published probes used in this study are alpha-1 globin/HBA1,15 alpha-2 globin/HBZP,16 and a globin RA330.9

Isolation and characterization of interspecific somatic cell hybrids. The somatic cell hybrid lines (H-101 and H-201) were made by the method of Deisseroth and Hendrick14 as modified by Zeiten15 and Weatherall.13 Epstein-Barr virus (EBV) transformed lymphocytes from the patient R.A. were fused with adenine phosphoribosyl transferase (APRT) negative mouse erythroleukemia (MEL) cells (line 585, a gift from A. Deisseroth, University of California), and hybrid cells containing human chromosome 16 were selected in medium containing metotrexate (0.1 mmol/L), adenosine (0.1 mmol/L), thymidine (30 mmol/L), and ouabain (0.5 mmol/L). Ouabain was only included during the first 14 days of culture to prevent the background growth of EBV-transformed lymphocytes in preference to the mouse/human hybrid cells. Mapping of the alpha genes and karyotyping of two independent clones showed that the normal (alpha) chromosome 16 was present in a tetraploid MEL background (line H-201), while the abnormal (alpha)RA chromosome 16 was present in a diploid MEL background (line H-101). In situ hybridization to the cell line H-101, performed as previously described using 3H-labeled total human DNA,18 demonstrated that each of the 50 cells examined contained only one human chromosome. To study globin gene expression, MEL and hybrid cell lines were induced with DMSO (1.4% vol/vol) and hemin (40 mmol/L) for 3 days before analysis of RNA and globin chain synthesis.

RNA analysis. Total RNA was prepared from cell pellets (usually 2 to 5 x 10^7 cells) that had been kept at -70°C until extraction, as described by Maniatis et al19 or by the method of Chomczynski and Sacchi.20 Human or mouse alpha globin mRNA was detected using the quantitative RNAase mapping procedure.21 Approximately 1 mg of a plasmid (psp601322) containing an insert corresponding to the 5' end of both the human alpha and beta genes was linearized with BamHI. Similarly, 1 mg of plasmid (psp6422) containing an insert corresponding to the 5' end of the mouse alpha globin gene was linearized with HindIII. A 3P-labeled probe was transcribed from each template using SP6 polymerase as specified in the SP6 system (Amersham International plc, UK) and [alpha-32P]GTP (410 Ci/mmol).22,23 Each probe (approximately 1 x 10^6 cpm) was added separately to 10 mg of total cellular RNA, heated to 95°C for 10 minutes, and then hybridized in 80% formamide, 40 mmol/L PIPES (pH 6.4), 400 mmol/L NaCl, and 1 mmol/L EDTA, at 50°C overnight. The resulting RNA-RNA hybrids were then treated with

From MRC Molecular Haematology Unit, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford; and Department of Haematology, Princess Margaret Hospital, Swindon, UK.

Submitted January 29, 1990; accepted March 12, 1990.

Address reprint requests to D.R. Higgs, DSc (Med), MRC Molecular Haematology Unit, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, OX3 9DU UK.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1990 by The American Society of Hematology.

0006-4971/90/7601-00733.00/0
Table 1. Hematologic Data in Patient R.A. and His Offspring

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age (y)</th>
<th>Hb (g/100 mL)</th>
<th>MCV (fL)</th>
<th>MCH (pg)</th>
<th>HbH Inclusions</th>
<th>α/β Globin Chain Synthesis Ratio</th>
<th>BamHI α-Specific Fragments (kb)*</th>
<th>BglII ψ′ 1-Specific Fragments (kb)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.A. M</td>
<td>56</td>
<td>13.2</td>
<td>64</td>
<td>22</td>
<td>+</td>
<td>0.71</td>
<td>14.0</td>
<td>12.6, 11.0, 6.8</td>
</tr>
<tr>
<td>R0.A. M ad</td>
<td>12.0</td>
<td>71</td>
<td>22</td>
<td>NS</td>
<td></td>
<td>0.66</td>
<td>14.0</td>
<td>12.6, 10.0, 6.8</td>
</tr>
<tr>
<td>D.A. F ad</td>
<td>12.8</td>
<td>72</td>
<td>24</td>
<td>NS</td>
<td></td>
<td>NS</td>
<td>14.0</td>
<td>12.6, 10.0, 6.8</td>
</tr>
</tbody>
</table>

The mother of R0.A. and D.A. was unavailable for study.

Abbreviations: ad, adult at time of study; NS, not studied; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin.

*The 14.0-kb fragment spans from +14 to +28 and contains both the α1 and α2 genes. This normal fragment is not shown in Fig 1.
†Only the hypervariable segment (10 to 11.0 kb) and the breakpoint fragment (6.8 kb) are shown in Fig 1.

The protected fragments (128 nucleotides for mouse α and 133 nucleotides for both human α1 and α2) were analyzed on an 8% denaturing polyacrylamide gel.

Molecular cloning of the α globin genes from the (αα)RA chromosome. High molecular weight DNA obtained from the hybrid cell line H-101, containing the abnormal (αα)RA chromosome, was digested to completion with BamHI. Unfractionated DNA was ligated using standard conditions to BamHI-cut bacteriophage arms prepared from the vector λEMBL 3, which had previously been treated with EcoR1 and alkaline phosphatase (Promega). Two micrograms of the ligation mix were packaged in vitro using a standard protocol (Lambda in vitro packaging kit, Amersham) and plated on the host bacterial strain NM621. Recombinants, 4 × 10⁵, were screened with the α1 globin probe as described in Kaiser and Murray, and one positive plaque (λ AW1) was identified. Subsequently, a bulk preparation of DNA from the recombinant phage was prepared. When analyzed by digestion with BamHI and blot hybridization to α globin, the predicted 14-kb band was identified. However, a 10.5-kb fragment, which varied in proportion between different preparations, was also present. This was thought to have arisen by spontaneous deletion as a result of recombination between the tandem α globin genes as previously described. The correct 14-kb fragment was isolated by bulk digestion of λ AW1 with BamHI, agarose gel electrophoresis, excision of the relevant band, and electroelution. The absence of the contaminating 10.5-kb band was confirmed and the presence of both α1 and α2 genes in the 14-kb BamHI band was demonstrated by Apal/Psil double digestion.

RNAases A (40 μg/mL) and T1 (2 μg/mL) at 16°C for 30 minutes.

Fig 1. (Top) The normal human α globin cluster (denoted αα) with reference points given in kilobases. Position 0 represents the 2 mRNA CAP site. Functional genes; ( ), pseudogenes, and ( ), the region corresponding to the probe α globin RA330. The solid bar beneath the normal cluster represents the extent of the deletion from the normal chromosome to produce the abnormal (αα)RA chromosome. The BamHI (B), BglII (Bg), and HindIII (H) restriction sites nearest to the breakpoints are indicated. (Middle) The abnormal chromosome, (αα)RA, is shown together with the abnormal BamHI, BglII, and HindIII fragments detected by the ψ′1 probe. (Insert) Examples of normal genomic DNA (1), DNA from a hybrid containing a normal chromosome 16 (2), a hybrid containing (αα)RA (3), and MEL cell DNA (4) hybridized to the ψ′1 probe. Abnormal fragments also identified by α globin RA330 are marked thus ( ). The relative migration of DNA molecular weight markers (HindIII cut λ) is shown on the right of the insert panel. Fragment sizes are 23, 9.6, 6.5, 4.3, 2.3, and 2.0 kb.
A DELETION UPSTREAM OF THE α GLOBIN CLUSTER

Subsequently, the 1.5-kb PstI band from the 14-kb fragment was subcloned into PstI cut and alkaline phosphatase-treated pBR322 (Pharmacia, Sweden). Recombinants containing the α1, α2, or ψα1 genes, each of which is contained within the 1.5-kb PstI band, were distinguished by digestion with Apal and RsaI, and those containing the α1 and α2 genes are referred to as pRAα1 and pRAα2, respectively. Finally, the α1- and α2-specific PstI fragments were each subcloned in both orientations into the PstI site of M13tg and are referred to as M13RAα1, M13RAα1(R), and M13RAα2, M13RAα2(R).

Sequence analysis of the α1 and α2 globin genes. 3' End-labeled Ncol/PstI and HindIII/PstI fragments were prepared from pRAα1

Table 2. Mapping Data for the Normal (αα) and Abnormal (αα)α Chromosomes

<table>
<thead>
<tr>
<th></th>
<th>$\psi^\alpha$</th>
<th>α Globin RA330</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>αα</td>
<td>αα</td>
</tr>
<tr>
<td>BamHI</td>
<td>10.8, 5.9</td>
<td>14.0</td>
</tr>
<tr>
<td>BglII</td>
<td>12.6, 11.0</td>
<td>12.6, 6.8</td>
</tr>
<tr>
<td>HindIII</td>
<td>16.0, 13.0</td>
<td>12.6, 1.5</td>
</tr>
<tr>
<td>EcoRI</td>
<td>23.0, 5.0</td>
<td>&gt;23</td>
</tr>
</tbody>
</table>

Fragments underlined are common to both the $\psi^\alpha$ and RA330 probes.

Fig 2. (A) Analysis of mouse α globin mRNA in total RNA from uninduced (−) and induced (+) hybrid cell lines H-101, H-201, MEL, and the human cell line K562, which has been previously shown to produce α globin mRNA. The relatively faint band in K562+ at 128 nucleotides represents a low level of cross contamination from the strong signal in the MEL+ lane, with no corresponding band having been seen on many other occasions. (B) Analysis of human α globin mRNA in total RNA from uninduced (−) and induced (+) hybrid cells H-101, H-201, MEL, and K562.
and pRAa2, and sequenced using the method of Maxam and Gilbert. In addition, single-stranded DNA prepared from M13Raa1, M13Raa1(R), and M13Raa2, M13Raa2(R) was sequenced by the dideoxy sequencing reaction as described in the protocol "DNA sequencing with Sequenase" (US Biochemical Corporation, 1988). Priming from oligonucleotides corresponding to sequences within and flanking the structural α genes allowed analysis of the entire sequence.

RESULTS

The 62-kb deletion is present on the same chromosome as the α-thalassemia determinant. The propositus (R.A.) was a 56-year-old man who initially complained of anorexia and weight loss for which no cause was found. The presenting symptoms eventually resolved without treatment. A routine blood count showed a hypochromic microcytic anemia in the absence of iron deficiency. Hb electrophoresis showed normal proportions of HbA2 (2.7%) and Hbf (1.1%). A diagnosis of α-thalassemia was confirmed by demonstrating HbH inclusions in the peripheral red blood cells (RBCs) and a reduced α/β globin chain synthesis ratio (Table 1). Blot hybridization studies using DNA obtained from the peripheral blood of R.A. demonstrated that there was a 62-kb deletion from upstream of the α globin complex on one chromosome (see reference 9 and Fig 1). To determine whether the 62-kb deletion is linked to the α-thalassemia determinant, we examined two offspring (Ro.A. and D.A.) of the propositus. Both have the phenotype of α-thalassemia and both have inherited the (αα)RA chromosome, identified by an abnormal (6.8 kb) β-specific BglII fragment (Table 1 and Fig 1). These findings suggest that the phenotype of α-thalassemia in this family can be entirely explained by the inheritance of the (αα)RA chromosome since the degrees of chain imbalance and the RBC indices of R.A. and his children, all of whom have the abnormal 6.8-kb BglII fragment, are nearly identical (Table 1).

Functional analysis of the somatic cell hybrids. To simplify the subsequent structural and functional analysis of the (αα)RA determinant, each chromosome 16 from the propositus was isolated in human × MEL somatic cell hybrids. The hybrid containing the (αα)RA chromosome was mapped with BamHI, BglII, and HindIII, and the expected breakpoint fragments were identified (see Table 2 and Fig 1). It has been previously demonstrated that expression of both the endogenous mouse globin genes (α, β major, and β minor) and the introduced human α genes (α2 and α1, but not β) can be induced by treating the hybrid cells with a variety of agents. When H-201 hybrids containing the normal αα chromosome were treated with hemin and DMSO, both mouse and human α globin mRNA were induced (Fig 2, A and B). Furthermore, both mouse and human globin chain synthesis could be readily detected after induction of these cells (Fig 3). By contrast, under the same experimental conditions only a trace of human α mRNA (less than 1% of mouse α globin mRNA) and no human α globin synthesis could be detected in the cell line H-101 containing the abnormal (αα)RA chromosome (Figs 2 and 3). Thus, these data confirm that the α-thalassemia determinant is present in cis to the deletion on the (αα)RA chromosome.

Sequence analysis of the α globin genes on the (αα)RA chromosome. Because the α thalassemia determinant lies in cis to the 62-kb deletion, either the α genes are inactivated as a direct result of the deletion or another inactivating mutation(s) must exist on this chromosome. The most likely site for such a mutation(s) would be within the transcription units of the α genes themselves. Therefore, both α1 and α2
genes were cloned from DNA obtained from H-101, which contains the (αα)RA chromosome, and the sequences from −175 to +897 (α1) or +893 (α2) with respect to the mRNA CAP site were determined (ie, from 175 nucleotides upstream of the transcription initiation site to 50 nucleotides downstream of the poly(A) addition site). The sequences of both the α1 and α2 genes between −175 and the poly(A) addition site (see Fig 4) were identical to previously published sequences from nonthalassemic individuals (see legend to Fig 4). However, in the α1 gene there were three previously undescribed variant nucleotides in the 3’ noncoding region beyond the poly (A) addition site (Fig 4), two of which have also been noted in the sequence of a functional α gene (Horst J. and Griese E.-U., personal communication, January 1990). In the α2 gene there was a previously undescribed variant nucleotide 27 base pairs beyond the poly (A) addition site (Fig 4), two of which have also been noted in the sequence of a functional α gene (Horst J. and Griese E.-U., personal communication, January 1990). Two other differences between these sequences and the published sequence of Michelson and Orkin were also noted. Double mutations affecting the β globin cluster have been described in Sardinian patients with an unusual form of ββ-thalassemia and in a single patient with a β-thalassemia mutation in cis to a β variant. Double mutations of the α genes have also been described in Algerian patients. In this case there is a large deletion (−α) and a small, dinucleotide deletion in the remaining α gene. However, such occurrences would be quite unexpected in an individual from a population in which α-thalassemia is otherwise rare.

We have now clearly demonstrated that the (αα)RA chromosome, with the 62-kb deletion, accounts for the α-thalassemia phenotype in this family. Although this suggests even more strongly that the deletion is primarily responsible for the α-thalassemia, it is clearly impossible to rule out a second inactivating mutation on this chromosome with absolute certainty, since not all of the cis-acting sequences required for α globin gene expression have yet been identified. However, no mutations were found within the transcription units of the α1 and α2 genes from the (αα)RA chromosome, and it seems very unlikely that the previously unreported sequence changes in their divergent 3’ noncoding regions could severely downregulate both α genes.

Thus, the most plausible explanation is that the deletion is primarily responsible for downregulating the expression of the α globin genes. This could be due to a negative element.

**DISCUSSION**

The initial observation of a large deletion upstream of the α globin genes in a patient with α-thalassemia suggested that this mutation was responsible for the phenotype. Nevertheless, we could not exclude the possibility that there was a second mutation either in cis or trans to this deletion that accounted for the α-thalassemia. Such instances have been noted in populations in which thalassemia is common. For example, deletions of the α cluster that have no effect on phenotype, such as those that result in a chromosome with a single Δ gene (ΔI) rather than the normal (Δ2-Δ3-1) rearrangement, may occur in cis or trans to other α-thalassemia mutations. Furthermore, double mutations affecting the β globin cluster have been described in Sardinian patients with an unusual form of ββ-thalassemia and in a single patient with a β-thalassemia mutation in cis to a β variant. Double mutations of the α genes have also been described in Algerian patients. In this case there is a large deletion (−α) and a small, dinucleotide deletion in the remaining α gene. However, such occurrences would be quite unexpected in an individual from a population in which α-thalassemia is otherwise rare.

We have now clearly demonstrated that the (αα)RA chromosome, with the 62-kb deletion, accounts for the α-thalassemia phenotype in this family. Although this suggests even more strongly that the deletion is primarily responsible for the α-thalassemia, it is clearly impossible to rule out a second inactivating mutation on this chromosome with absolute certainty, since not all of the cis-acting sequences required for α globin gene expression have yet been identified. However, no mutations were found within the transcription units of the α1 and α2 genes from the (αα)RA chromosome, and it seems very unlikely that the previously unreported sequence changes in their divergent 3’ noncoding regions could severely downregulate both α genes.

Thus, the most plausible explanation is that the deletion is primarily responsible for downregulating the expression of the α globin genes. This could be due to a negative element.
from beyond coordinate — 52 being juxtaposed close to the α genes. Alternatively, the large 62-kb deletion could perturb the higher order chromatin structure around the α globin complex in a relatively nonspecific manner, although other large insertions (+ 10 kb) and deletions (− 10 kb) within the α complex do not appear to alter the expression of the α genes in a significant way.8,39-41

The third possible mechanism by which the deletion could inactivate the α genes is by the removal of a specific positive regulatory element(s) that is essential for their expression. The existence of such positive regulatory sequences is suggested by previous observations that when DNA fragments containing the α or β globin genes are assayed in experimental erythroid systems (transient assays,44,45 stable transformants,46 or transgenic mice,46,47 the genes are either inactive or, at best, are expressed at a level that is considerably less (often less than 1%) than that of the endogenous genes. Furthermore, the levels of expression are dependent on the site of integration in the genome. However, α and β genes transferred to interspecific human × MEL cell hybrids on chromosomes 16 and 11, respectively, are induced and expressed at a level approximately equal to that of the endogenous mouse globin genes.17,34,45,48 The implications of these observations are that additional cis-regulatory sequences, remote from the structural genes, are required to produce high levels of tissue-specific expression that are independent of the position of integration.

It has recently been demonstrated that such positive regulatory sequences exist upstream of the human β globin gene cluster.49 Furthermore, three different deletions that remove these sequences severely downregulate β globin gene expression.50-52 The positive regulatory sequences in the β globin cluster (referred to as the β-dominant control region[β-DCR] or β-locus activating region [β-LAR]) confer high level and position independent expression on the β-like genes when constructs containing both the DCR and the gene(s) are transfected and integrated into the genome of MEL cells or transgenic mice.49,53-56 Recently, it has also been shown that the β-DCR can exert a similar effect on the α globin genes in these experimental systems.56,47 Therefore, it seems probable that similar positive regulatory sequences exist upstream of the human α globin cluster and that these sequences are deleted in the (aa)8 chromosome. Therefore, the characterization of this naturally occurring mutant points to the region to be investigated in the search for such positive regulatory sequences in the α globin complex.

ACKNOWLEDGMENT

We are grateful to D.J. Weatherall for his encouragement and support. We also thank R.A. and his family for their help and cooperation in this study. We thank Margaret Baron for kindly supplying the plasmids used in the nucleotide protection assays. We are grateful to Professor J. Horst and Dr E.-U. Griese for communicating their unpublished sequence data from the — α17 determinant.

REFERENCES

21. Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K,
A DELETION UPSTREAM OF THE α GLOBIN CLUSTER


37. Morle F, Lopez B, Henri T, Godet J: α-Thalassemia associated with the deletion of two nucleotides at position −2 and −3 preceding the AUG codon. EMBO J 4:1245, 1985


Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human alpha globin gene cluster

CS Hatton, AO Wilkie, HC Drysdale, WG Wood, MA Vickers, J Sharpe, H Ayyub, IM Pretorius, VJ Buckle and DR Higgs