Volume 12 Number 18 1984 Nucleic Acids Research

Independent recombination events between the duplicated human α globin genes; implications for their concerted evolution

D.R.Higgs, A.V.S.Hill, D.K.Bowden+, D.J.Weatherall and J.B.Clegg

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Level 7, Headington, Oxford OX3 9DU, UK, and +Department of Anatomy, Monash University, Clayton, Victoria, Australia

Received 19 July 1984; Revised and Accepted 4 September 1984

SUMMARY

We have examined the molecular structure of the human α globin gene complex from individuals with a common form of α thalassaemia in which one of the duplicated pair of α genes (αα) has been deleted ($-\alpha^{3.7}$). Restriction mapping and DNA sequence analysis of the mutants indicate that different $-\alpha^{3.7}$ chromosomes are the result of at least three independent events. In each case the genetic crossover has occurred within a region of complete homology between the α1 and α2 genes. Since the −α chromosomes may reflect the processes of crossover fixation and gene conversion between the two genes, their structures may provide some insight into the mechanism by which the concerted evolution of the human α globin genes occurs.

INTRODUCTION

The human α globin gene complex on chromosome 16 includes two α genes (α1 and α2), an embryonic α-like gene (ζ) and two pseudogenes (ψα and ψζ) within a 30kb segment of DNA, arranged in the order 5'-ζ-ψζ-ψα-α2-α1-3' (1,2). This multigene family is thought to have evolved by a series of gene duplications followed by sequence divergence punctuated by insertions and deletions of DNA, surrounding the coding portions of the genes (2-6). Heteroduplex and DNA sequence analysis has demonstrated that the α globin genes are embedded within two homologous segments of DNA, each approximately 4kb long (2,6). Insertions or deletions have further divided each of these regions into three homologous subsegments (X, Y and Z) (Figure 1), the Zα1 and Zα2 segments containing the α1 and α2 globin genes respectively (2,6). Interspecies comparisons indicate that the ancestral α globin gene was duplicated at least 270 million years ago (7) and while the sequences of the two human α genes have diverged from other species as predicted, they still direct the synthesis of identical peptide chains (8). Furthermore, comparison of the entire DNA sequences of the Zα1 and Zα2 segments (Figure 1) shows that they are highly conserved (6). This type of evolution-in-tandem is called concerted evolution and two mechanisms have been proposed to explain this phenomenon, gene conversion and crossover fixation (Figure 2). There is evidence to suggest that both processes may be involved

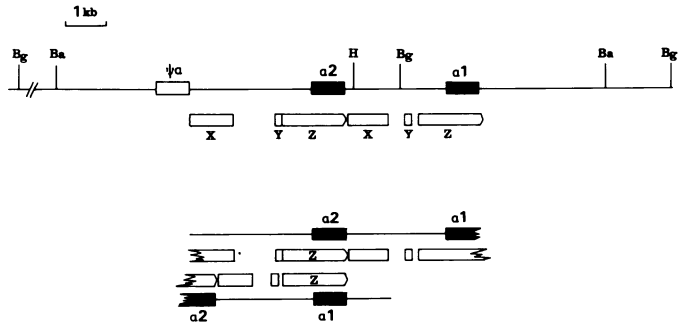

Figure 1: **Above**: Details of the α globin duplication unit showing the highly conserved X, Y and Z homology blocks. Restriction sites Bg (Bgl II), Ba (Bam H1) and H (Hpa I) are shown. **Below**: Putative arrangement of the misaligned Zα1 and Zα2 regions which allows exchange of genetic material between these two highly homologous segments of DNA.

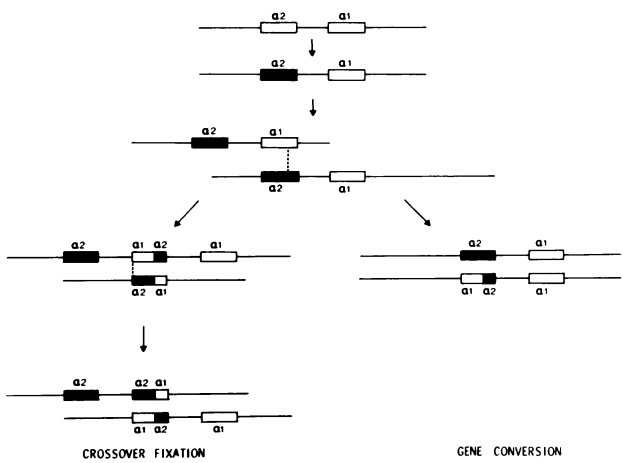

Figure 2: Hypothetical models to explain the concerted evolution of the human α globin genes (summarised from 2,4,6 and 7). Following duplication of the ancestral α gene to produce α1 and α2 genes, the sequences diverge during evolution as indicated. Misalignment occurs, which allows exchange of genetic material between Zα1 and Zα2 as outlined in Figure 1. In one case (right) a non reciprocal exchange (gene conversion) occurs giving rise to one chromosome in which part of the divergent α2 sequence has been corrected back to the α1 sequence with no change in gene number. In the other model (left), two rounds of reciprocal crossover have taken place (crossover points marked by dashed lines). In this case chromosomes with single or triplicated α genes resulting from reciprocal crossovers occur as intermediates in the cycle (7). Again the outcome is one chromosome in which the α2 has been partially corrected back to the α1 sequence. If these new (α1/α2, α1) chromosomes become fixed in the population, the sequence divergence between α1 and α2 in the region of the gene correction will no longer be apparent in that population.

in maintaining the identity of the human α globin genes and their flanking DNA, although their relative importance, frequency, and precise underlying mechanisms are not known.

If the DNA sequences of these regions have been kept identical by crossover fixation, it would be predicted that intermediates in the cycle of gene duplication and loss should exist within a population. The products of reciprocal recombination between homologous Z segments, which are 3.7 kb apart, have been observed in vivo and in vitro; these give rise to chromosomes with single ($-\alpha^{3.7}$) or triplicated ($\alpha\alpha\alpha^{anti3.7}$) α genes and the $-\alpha^{3.7}$ haplotype is a common cause of α^{+} thalassaemia in many populations (reviewed in reference 9). The finding of chromosomes with one (−α) and three (ααα) genes within different racial groups therefore supports crossover fixation as a mechanism for maintaining sequence identity between the two genes. However, since for all organisms where genetic analysis has been possible, conversion appears to be non-randomly associated with reciprocal recombination (10), these observations are also consistent with gene conversion being the driving force behind the concerted evolution of the α genes.

If the entire sequence of the Z boxes has been maintained by these mechanisms, the same degree of homology between Zα1 and Zα2 should exist throughout the entire Z regions. In fact the homology between these two regions is segmental, two adjacent tracts of DNA being >99% and 93% homologous respectively (Figure 3). The latter region may itself be subdivided into 3 regions which are >99%, 78% and 100% homologous (Figure 3). To account for the two regions that are >99% and 93% homologous it has been suggested that the Z segments comprise two major gene conversion units separated by a 7 bp insert in IVS 2 of the α1 gene (△ in Figure 3) which may behave as a barrier to gene conversion by inhibiting branch migration following the formation of an α1/α2 heteroduplex (6). Since this insert divides the Z box into long (1436 bp) and short (282 bp) regions it is possible that the larger target area presented by the region 5′ to the insert allows more opportunity for the exchange of genetic material in this segment than in the adjoining region of DNA and hence, through more frequent gene conversion events, is more highly conserved than the adjacent less converted region. A prediction of this model is that in the naturally occurring $-\alpha^{3.7}$ defects the points of crossover should mainly lie on the 5′ side of the 7 bp insert. In keeping with this hypothesis, previous analysis of sixteen $-\alpha^{3.7}$ chromosomes from different populations showed that in all cases, crossover had occurred before the 7 bp insert, and in one such crossover the $-\alpha^{3.7}$ chromosome appeared to be a

"patchwork" sequence resulting from a reciprocal crossover in an α1/α2 heteroduplex; the remaining Z box contained single nucleotides which were both α1- and α2-specific, providing good indirect evidence for an associated gene conversion event (6).

Even this modified model does not explain why the homology within the region beyond the 7 bp insert is also segmental. Therefore, in view of this complicated pattern of segmental homology we have examined further examples of the $-\alpha^{3.7}$ defect in the belief that they might represent intermediates in the processes involved in the concerted evolution of the α globin genes and may thus provide information on the mechanism by which this occurs. We have identified chromosomes resulting from at least three independent crossover events, and, surprisingly, in two of them the points of resolution lie within the homologous regions located downstream from the 7 bp insert. These findings may indicate that the recombination event in these cases was initiated and terminated beyond this region. Alternatively, the 7 bp insert may not always act as a barrier to gene conversion. Finally, we have examined the region of the previously reported "patchwork" structure and show that this sequence is polymorphic in several populations and cannot therefore be used to specify the events producing the $-\alpha^{3.7}$ chromosome.

MATERIALS AND METHODS

Heparinised peripheral blood was obtained from α thalassaemic and non-thalassaemic individuals of Jamaican, Southeast Asian, Mediterranean and Melanesian origins. Haematological studies and haemoglobin analyses were performed using standard techniques (11). DNA was prepared from the buffy coats of each sample by phenol-chloroform extraction (summarised in reference 12). Blot hybridisation studies using various restriction enzymes and a ^{32}P labelled α-specific probe (a 1.5 kb PstI fragment from the genomic α subclone pKBα1 (2)) were carried out as previously described (12). $-\alpha^{3.7}$ deletion chromosomes were identified by the presence of a 10.5 kb Bam H1 and 16 kb Bgl II α-specific bands (2).

Cytoplasmic mRNA was prepared from the peripheral blood reticulocytes of some patients (13) and the type of α-specific mRNA present (α1 or α2) was determined by single strand nuclease analysis essentially as described by Orkin and Goff (14).

In one α thalassaemic individual, homozygous for the Hb variant J Tongariki ($\alpha^{115\ Ala->Asp}$) with the genotype $-\alpha^J/-\alpha^J$ (15), a recombinant containing a single α gene ($-\alpha^J$) was isolated from a library of size

fractionated, EcoRI cut DNA cloned into the λ phage substitution vector L47.1 (16,17). Subsequently, a Hind III - EcoRI fragment containing the 3' half of the α^J gene and approximately 2 kb of the 3' non-coding and flanking regions was subcloned into pBR322 (pDH10). Fragments from this recombinant were 3' end labelled (17), degraded and the products were analysed on 6% denaturing gels (18).

RESULTS

(i) The $-\alpha^{3.7}$ defects have different 3' non-coding regions

In a survey of individuals with α thalassaemia from several different populations, we identified 293 chromosomes containing the $-\alpha^{3.7}$ defect. In order to determine in a quick and simple way whether the structures of these $-\alpha^{3.7}$ defects were heterogeneous, we initially performed S1 nuclease analysis of the mRNA expressed by the genes. There is sufficient divergence in the 3' non-coding tracts of the normal α globin genes to allow the type (α1 or α2) and relative quantity of mRNA expressed to be determined in α^+ thalassaemia (14,19). Hitherto, analysis of mRNA in all examples of $-\alpha^{3.7}$ defects has shown that it is of the α1 type and therefore that the crossovers which generated them have occurred 5' to the divergent regions of DNA in the 3' non-coding regions. Analysis of mRNA from nine individuals of Jamaican, Mediterranean, North European and Melanesian origins demonstrated that in some of the Melanesian $-\alpha^{3.7}$ haplotypes the 3' end of the mRNA was α2 - rather than α1- like showing that the crossover responsible for this defect had occurred within or beyond the region of divergence (figure 3). In particular, the $-\alpha^{3.7}$ haplotype in which there is an Ala-Asp mutation at position 115 (Hb J Tongariki) invariably produced mRNA of the α2 type.

(ii) Localisation of the recombination sites

Since mRNA analysis showed that the group of α^+ thalassaemias which result from homologous recombination events within the Z boxes ($-\alpha^{3.7}$) clearly result from at least two types of crossover, we analysed these defects in greater detail in order to further localise the exact sites of recombination. Fortunately, several of the differences in nucleotide sequence between the α1 and α2 genes can be identified by restriction enzyme analysis. For example, it has previously been shown that the enzyme ApaI recognises several sites of sequence divergence between the α1 and α2 regions of DNA (Figure 4) (6). Similarly, we have found that the enzymes Bal I and Rsa I also recognise several differences in the nucleotide sequences

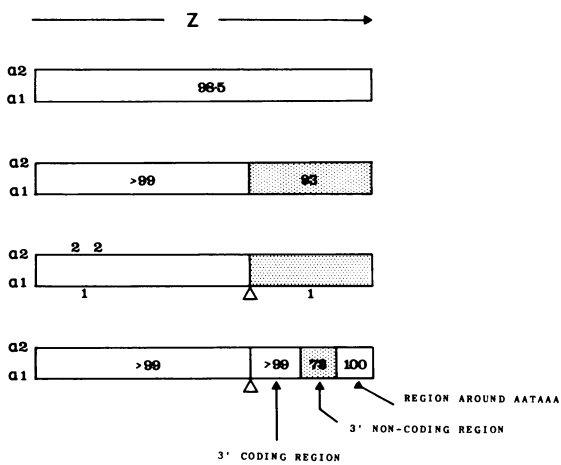

Figure 3: Diagramatic representation of the % homologies between the human Zα1 and Zα2 segments containing the α1 and α2 genes respectively (not to scale). While the overall homology is 98.5% this may be divided by a 7 bp insert in IVS 2 of the α1 globin gene (6) into regions of >99% and 93% homology. It has been suggested that these two segments represent long and short conversion units. A prediction of this model is that crossovers and conversions should occur more frequently within the long homology unit. Hitherto this prediction has been substantiated and in one case of an $-\alpha^{3.7}$ defect the sequence indicated that it was an α2/α1/α2/α1 patchwork gene (denoted by 2,1,2,1 in the diagram) supporting the concept that it resulted from a gene conversion/crossover event. In contrast, the data presented here demonstrate that the apparent patchwork sequence in the 5' region of the Z box results from the presence of a Rsa I restriction site polymorphism, and that crossovers do occur in the two homologous subsegments beyond the 7 bp insert. Finally, the homology in the region beyond the 7 bp insert is also segmental as shown.

between the two genes (Figure 4 and 5). The combined results of blot hybridisation studies using these three enzymes on DNA from the $-\alpha^{3.7}$ defect enables the point of crossover to be localised to one of several segments of DNA within the Z box (Table 1 and Figures 4,5 and 6). We observed patterns of α-specific fragments consistent with crossovers in three of these regions (labelled I, II and III in figures 4,5, and 6). Of the total $-\alpha^{3.7}$ defects examined, crossovers in segment I which were indistinguishable from those previously reported (6) were observed in all populations studied. In addition, we observed the products of hitherto undescribed recombination events in segment II (Jamaican, Southeast Asian and Mediterranean) and segment III (Melanesian and Polynesian). In chromosomes with the $-\alpha^{3.7}$ defect arising from crossovers within segments I and II (Figure 4) the points of recombination clearly lie within regions of complete homology between the α1

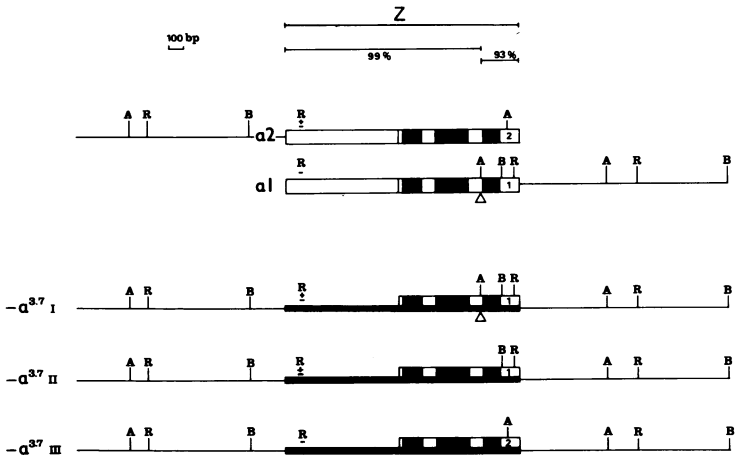

Figure 4: **Above**: Misaligned Zα1 and Zα2 segments showing the relative positions of A (Apa I), B (Bal I) and R (Rsa I) restriction sites. ± indicates the presence or absence of a polymorphic Rsa I restriction site at the 5′ end of the Z box. Each Z box includes the 5′ flanking region ▭, 5′ non-coding region ▭, coding regions ▬, intervening sequences □ and divergent 3′ non-coding regions [1] and [2]. △ indicates the 7 bp insert in IVS 2 of the α1 gene. **Below**: Restriction maps of the $-\alpha^{3.7}$ defects resulting from crossovers in I before the 7 bp insert in α1, II between the Apa I and Bal I sites in α1, III beyond the Apa I site in α2 and Rsa I site in α1.

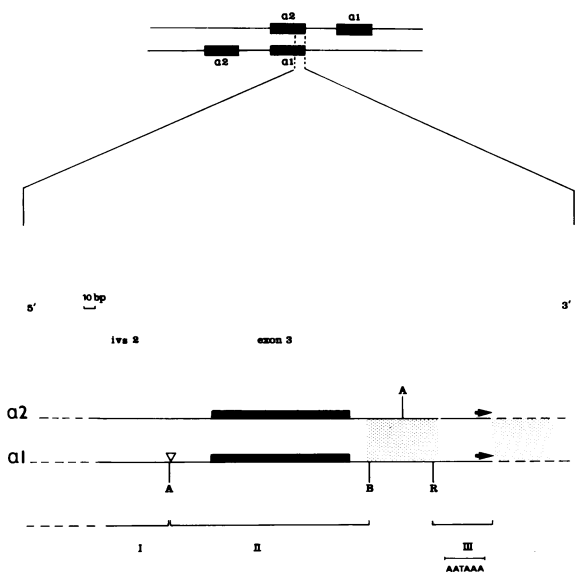

Figure 5: **Above** The α1 and α2 globin genes in the positions assumed during a misaligned crossover. **Below**: Detailed comparison of the α1 and α2 genes beyond the 7 bp insert (△) in α1. Restriction sites A (Apa I), B (Bal I) and R (Rsa I) are indicated. ➡ indicates the sequence 5′-CCTG(TG)$_3$CCTG-3′ which marks the end of the human α globin duplication unit. Segments I, II and III correspond to the $-\alpha^{3.7}$ crossover regions indicated in Figure 4.

TABLE I. α-specific restriction fragments obtained in the αα and $-\alpha^{3.7}$ chromosomes.

	SEGMENT I	SEGMENT II	SEGMENT III
	αα	-α	-α	-α
Apa I	2.7, 1.7, 0.89	2.5, 0.89	3.5	2.7, 0.7
Bal I	3.56, 1.95, 1.6	1.80, 1.6	1.80, 1.6	3.4
Rsa I	3.5*, 1.75, 0.9	2.6*, 0.9	2.6*, 0.9	3.6

* this fragment may be 1.1 kb shorter in some individuals due to the presence of a Rsa I restriction site polymorphism as described in the text and as seen in subject 10 in figure 6.

and α2 genes. Furthermore, in those cases where restriction mapping indicated that the crossover had occurred in segments I and II the α mRNA was α1-like and for those in segment III the α mRNA was of the α2 type, as

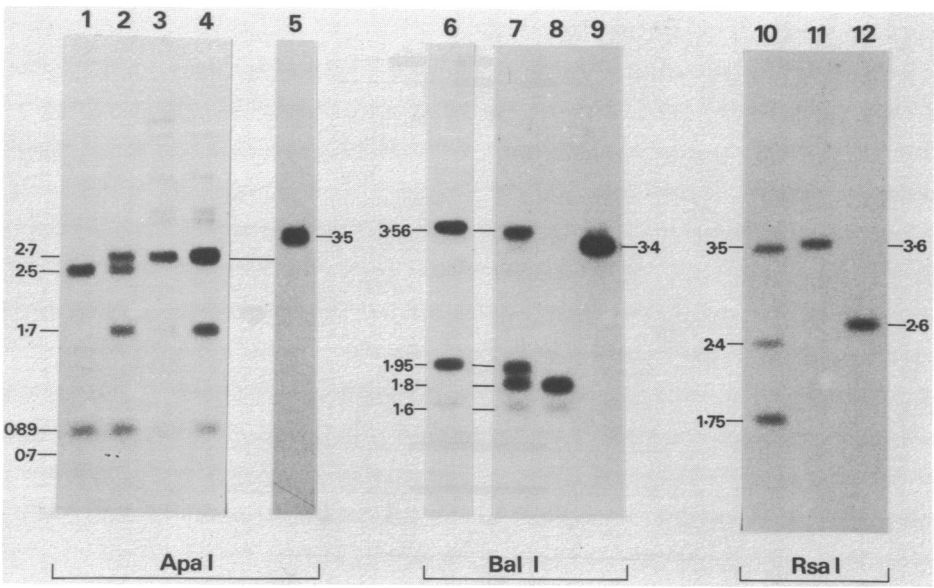

Figure 6: α-specific fragments obtained using Apa I, Bal I and Rsa I in subjects with the following genotypes (1) --/$-\alpha^{3.7}$ I (2) $-\alpha^{3.7}$ I/αα (3) $-\alpha^{3.7}$ III/$-\alpha^{3.7}$ III (4) αα/$-\alpha^{3.7}$ III (5) --/$-\alpha^{3.7}$ II (6) αα/αα (7) $-\alpha^{3.7}$ I/αα (8) $-\alpha^{3.7}$ I/$-\alpha^{3.7}$ II (9) $-\alpha^{3.7}$ III/$-\alpha^{3.7}$ III (10) αα/αα, one chromosome − and one + for the Rsa I polymorphism (11) $-\alpha^{3.7}$ III/$-\alpha^{3.7}$ III (12) $-\alpha^{3.7}$ I /$-\alpha^{3.7}$ I

expected. This latter pattern was observed in 80% of the Melanesian $-\alpha^{3.7}$ defects including all of those carrying the J Tongariki mutation.

(iii) Sequence analysis of $-\alpha^{3.7}$ J Tongariki

The precise location of the crossover points within the Z box are of particular interest with respect to whether they occur within regions of homology or non-homology. In the $-\alpha^{3.7}$ J Tongariki chromosome, restriction mapping indicated that the point of crossover was beyond the Rsa I site in α1 but before an Hpa I site which lies ~0.4 kb 3' to the end of the α2 gene (Figure 1) and could therefore lie in a region of homology or non-homology. In order to localise the recombination site, cloned DNA from the $-\alpha^{3.7}$ J Tongariki homozygote was sequenced across this region. This showed that the point of crossover lies within the 46 bp stretch of homology around the AATAAA sequence (segment III in Figure 6). Before this region the sequence is entirely α2-like, and beyond is entirely α1-like.

(iv) A polymorphic Rsa I site in the Z region

Although the majority of crossovers were located in segment I of the Z boxes, since these regions differ in only five nucleotides it is not possible to localise the crossover point precisely by restriction enzyme analysis. However, previous DNA sequence analysis of a single example of a $-\alpha^{3.7}$ defect in a Chinese individual (6) showed that this chromosome resulted from a crossover in segment I of the Z box and that the 5' region contained an α1/α2 "patchwork" nucleotide sequence identified by the presence of single nucleotide differences which are either α1- or α2-specific. At -869 the sequence of the $-\alpha^{3.7}$ defect was α2-like, at -733 α1-like, and at -634 α2-like again (6). This patchwork could have arisen in a conversion event in which there was no strand preference for correction. Inspection of the nucleotide sequence at -733 in the normal α1 and α2 Z segments shows that the α2 sequence (CAGTACC) includes a RsaI restriction site (GTAC) which is not present in the α1 sequence in the corresponding region (CAGCACC) (Figure 2). To determine whether the "patchwork" structure was present in all $-\alpha^{3.7}$ defects we analysed RsaI α-specific fragments in the $-\alpha^{3.7}$ chromosomes, which showed that it was present in some cases but absent in others. However, subsequent analysis of non-thalassaemic chromosomes (αα) from the same populations showed that the RsaI site is in fact polymorphic even within the normal (αα/αα) population. Therefore, it seems most likely that the original interpretation, that this region represents a "patchwork", was incorrect. Rather, this deletion resulted from a crossover involving a chromosome in which this polymorphic Rsa I site was absent in contrast to the normal sequence with which it was compared.

DISCUSSION

The co-evolution of a pair of active genes may result from a balance between gene conversion, crossover fixation and constraints on sequence divergence imposed by function. In this repect, it is interesting that in general, units of gene conversion are often limited to regions of DNA which contain functionally important sequences. In the particular case of the duplicated goat α globin genes for example, the conversion unit is only about 900 bp long and encompasses all known signals for accurate expression of the genes; the 5' boundary is formed by the CCAAT box and the 3' boundary by the region of the poly (A) addition site (20). Furthermore, in this case, the sequence of the two α genes is highly conserved throughout the entire conversion unit.

In contrast to these observations it is interesting to note the marked segmental nature of the homology between the human Zα1 and Zα2 regions which contain the recognised sequences for gene expression (6). This could be explained if the α genes undergo segmental gene conversion events limited by the ends of the α gene duplication unit (which include small poly purine-pyrimidine tracts of the type implicated in the initiation of recombination in prokaryotes and eukaryotes (21,22)) and a 7 bp insert in IVS 2 of the α1 gene (6). A major prediction of this model is that crossover events within the Z segment should overwhelmingly lie on the 5' side of the 7 bp insert. This report documents at least three independent crossover events and surprisingly in two of these the recombination site is beyond the putative barrier to genetic exchange (Segments II and III in Figures 4 and 5). Interestingly, the frequency and population distribution of these $\alpha^{3.7}$ defects is quite different (A.V.S. Hill et al, in preparation). In most populations, for example, we have identified crossovers only in regions I and II whereas in Melanesia the crossover in region III is the most prevalent $-\alpha^{3.7}$ defect. Superficially, these observations could be interpreted as a reflection of the rate at which the various crossovers occur in the large (Segment I, 1436 bp) medium (Segment II, 171 bp) and small (Segment III, 46 bp) regions of homology between the two α genes. Unfortunately, the putative selective advantage of α thalassaemia, and the possibility of genetic drift, does not allow these data to be used in assessing the relative frequency of crossover in these three segments of the Z box.

The existence of only two conversion units would be consistent with the observed differences in the 3' non coding regions of the α1 and α2 genes. However it does not adequately explain the lack of sequence divergence between

the two genes in their third exons and in the 46 bp regions surrounding their poly (A) addition signals (figures 3 and 5) while the segment of DNA joining these two regions has diverged considerably. It has been suggested that there could be functional constraints against synonymous codon changes in the third exon of the α genes in keeping with other observations which have shown that codon usage in active genes may be non-random (23). Alternatively, there could have been a "recent" gene conversion event limited to this block (Segment II in figure 5); segmental conversions have been proposed to explain the similar pattern of sequence homology between the human γ globin genes (24,25). Certainly, the observation of reciprocal crossover events in this region confirms that the DNA strand exchange necessary for gene conversions to occur do take place in this region, although it is not clear why such a conversion would not have also homogenised the remainder of the 3' non coding region since there is no obvious barrier to branch migration here. In addition, recent sequence data show that both human and both chimpanzee α genes are identical in their third exons (6,26), which would not be predicted if the sequences had been corrected "recently" following interspecies divergence, particularly since our preliminary restriction mapping indicates that the 7 bp insert is also present in the chimpanzee α1 gene where presumably it would also behave as a barrier to gene conversion. A third explanation would be that this small insert in IVS 2 does not act as a barrier to conversion and that branch migration initiated in the 5' Z segment may extend through the entire gene to just beyond the stop codon. Against this, but not exluding the possibility, lies the observation that to date all humans that we have studied have an α1- and α2- like gene with respect to IVS II; gene conversions extending through this region would be expected to have homogenised the two genes in this respect.

The same interpretations can be applied to the conserved 46 bp regions surrounding the poly (A) site of the two α genes. Small segmental gene conversion events have been documented in the human γ globin, immunoglobulin and mouse major histocompatability genes (24,25,27,28), and our observation of a reciprocal crossover in this region is consistent with this as a mechanism. However, as before it is difficult to understand why such a gene conversion event would not propagate into the non-homologous region of the 3' non-coding region. It has been previously noted that the sequence around the poly (A) site is not only conserved between the human α1 and α2 genes but also between human, rabbit and mouse α globin genes (29). Currently available sequence data allows us to extend these observations to chimpanzee (26), horse

(30) and goat (20). Whereas much of the 3' non-coding region has diverged as expected between these species, a 30 bp region including the poly(A) addition site has been conserved to a similar degree (77%-100%) to that seen in the 3rd exon of the α gene of these species (84%-100%). Hence we suggest that functional constraints must have played an important role in conserving this region. If these sequences had been conserved solely by gene conversion we would have expected to observe intra-species but not inter-species identity. Although the sequences surrounding the poly (A) signal are conserved in the α genes of different species this is not so when these sequences are compared with similarly located regions of the non-α globin genes (19). However, it has been previously noted that the 3' regions of the β globin genes (31) are also conserved between primate species suggesting that if these regions are functionally important they may play a specific functional role with respect to the gene they flank.

In conclusion, we have shown that the common $-\alpha^{3.7}$ deletion has been produced by at least 3 different crossover events. It is not possible to determine whether or not these represent intermediates in segmental gene conversion events which serve to maintain sequence homology between the two α genes during evolution. Interspecies comparison suggests that some of the observed homology in the 3' non-coding regions may have been maintained by functional constraints and in some cases these deletions may merely reflect the homology between the two genes which allow misalignment and reciprocal crossover to occur.

Acknowledgements

We are grateful to J. Lamb, H. Ayyub and H. Teal for excellent technical assistance. We are also very grateful to Professor G.R. Serjeant, P. Winichagoon, T. Sophocleous, Dr. R.C. Drysdale, Dr. R.J. Trent, Dr. S. Oppenheimer, and Dr. S.L. Thein for providing samples for this study. A.V.S. Hill holds an M.R.C. training fellowship. This work was supported in part by the Rockefeller Foundation (RF 83067).

REFERENCES

1) Diesseroth, A., Nienhuis, A., Turner, P., Velez, R., French Anderson, W., Ruddle, F., Lawrence, J., Creagan, R., Kucherlapati, R. (1977). Cell, 12:205-218.
2) Lauer, J., Shen, C-KJ, Maniatis, T. (1980). Cell, 20:119-130.
3) Proudfoot, N.J., Gil, A., Maniatis, T. (1982). Cell, 31:553-563.
4) Hess, J.F., Fox, M., Schmid, C., Shen, C-K.J. (1983). Proc. Natl. Acad. Sci., USA.
5) Proudfoot, N.J., Maniatis, T. (1980). Cell, 21:537-544.
6) Michelson, A. and Orkin, S.H. (1983). J. Biol. Chem., 258:15245-15254.
7) Zimmer, E.A., Martin, S.L., Beverely, S.M., Kan, Y.W., Wilson, A.C. (1980). Proc. Natl. Acad. Sci., USA, 77:2158-2162.

8) Foldi, J., Cohen-Solal, M., Valentin, C., Blouquit, Y., Hollan, S.R. and Rosa, J. (1980). Eur. J. Biochem., 109:463-470.
9) Higgs, D.R., Weatherall, D.J. (1983). Current Topics in Hematology. Alan R. Liss, NY. 4:37-99.
10) Fogel, S. and Hurst, D.D. (1967). Genetics, 57:455-481.
11) Dacie, J.V. and Lewis, S.M. (1975). In: Practical Haematology. 5th Ed. London:Churchill.
12) Old, J.M. and Higgs, D.R. (1980). In: The Thalassaemias, Ed. Weatherall, D.J. Churchill-Livingstone, Edinburgh, 74-102.
13) Hunt, D.M., Higgs, D.R., Old, J.M., Clegg, J.B., Weatherall, D.J., Marsh, G.W. (1980). Br. J. Haematol., 45:53-64.
14) Orkin, S.H. and Goff, S.C. (1981). Cell, 24:345-351.
15) Bowden, D.K., Pressley, L., Higgs, D.R., Clegg, J.B., Weatherall, D.J. (1982). Br. J. Haematol., 55:243-249.
16) Loenen, W.A. and Brammar, W.J. (1980). Gene, 10:249-259
17) Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). In: Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, New York.
18) Maxam, A.M. and Gilbert, W. (1980) Meth. Enzym., 65:499-560.
19) Liebhaber, S.A., Goossens, M., Kan, Y.W. (1981). Nature, 290:26-29.
20) Schon, E.A., Wernke, S.M. and Lingrel, J.B. (1982). J. Biol. Chem., 257:6825-6835.
21) Klysik, J., Stirdivant, S.M. and Wells, R.D. (1982). Nature, 299, 312-316.
22) Stringer, J.R. (1982). Nature, 296:363-366.
23) Miyata, T. and Hayashida, H. (1981). Proc. Natl. Acad. Sci., USA, 78:5739-5743.
24) Slightom, J.L., Blechl, A.E., Smithies, O. (1980). Cell, 21:627-638.
25) Stoechkert, C.J., Collins, F.S. and Weissman, S.M. (1984). Nucl. Acids Res., 12:4469-4479.
26) Liebhaber, S.A. and Begley, K.A. (1983). Nucl. Acids Res., 11:8915-8929.
27) Ollo, R. and Rougeon, F. (1983). Cell, 32:515;
28) Mellor, A.L., Weiss, E.H., Ramachandran, K. and Flavell, R.A. (1983). Nature, 306:792-795.
29) Nishioka, Y. and Leder, P. (1979). Cell, 18:875-852.
30) Clegg, J.B., Goodbourn, S.E.Y. and Braend, M. (Submitted for publication).
31) Martin, S.L., Zimmer, E.A., Davidson, W.S., Wilson, A.C. and Kan, Y.W. (1981). Cell, 25:737-741.