A novel rearrangement of the human β -like globin gene cluster

R.J.Trent, D.K.Bowden*, J.M.Old, J.S.Wainscoat, J.B.Clegg and D.J.Weatherall

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK, and *Vila Base Hospital, Vila Vila, Vanuatu

Received 28 October 1981

ABSTRACT

The first example of a duplication involving the human β -like globin genes has been characterised in DNA from a native of Vanuatu. Restriction endonuclease mapping has shown that a 5 kb insert of DNA in the γ - δ - β gene cluster is due to duplication of the γ -globin gene and results in a new rearrangement 5'- ϵ -G γ -G γ -A γ - δ - β -3'.

INTRODUCTION

In man the ϵ , γ , δ and β globin genes that code for the embryonic, fetal and adult β -like globin chains of haemoglobin are closely linked on chromosome 11 in the order 5'- ϵ -G γ -A γ - δ - β -3' (1,2,3). At birth there is a switch from fetal to adult haemoglobin production which is closely related to the gestational age of the fetus. The mechanism for the neonatal suppression of γ chains is not understood, nor is it clear why disorders such as HPFH or $\delta\beta$ thalassaemia, many of which are caused by gene deletions, result in persistent γ chain production into adult life (4).

This study presents the first example of a gene rearrangement in the β -like globin gene cluster in which an insertion of DNA rather than a deletion has occurred with a resultant duplication of the G γ gene to give a new rearrangement 5'- ϵ -G γ -G γ -A γ - δ - β -3'. Despite this rearrangement, the expression of the γ genes on the altered chromosome is only marginally affected.

MATERIALS AND METHODS

Patients The proband (SG) a 22 year old female from the island of Espiritu Santo in Vanuatu (New Hebrides) first

presented for a routine antenatal visit, at which time an elevated level in HbF was detected. Blood was subsequently obtained for DNA mapping from this patient, several members of her family and other unrelated individuals from the same population.

Haematological data and haemoglobin analysis Routine haematological analysis was carried out using a Coulter model S cell counter. Haemoglobin A₂ levels were ascertained by cellulose acetate electrophoresis and HbF levels were determined by a modification of the Betke method (4). The relative proportion of glycine residues at position γ 136 was determined by amino acid analysis of peptide γ CB3 which was isolated from a cyanogen bromide digest of unfractionated globin from SG (5). Gene mapping Buffy coats obtained from heparinised blood were stored at -90°C until DNA was extracted (6). 10 μg aliquots of DNA were digested with various restriction enzymes and fractionated on vertical 0.8% agarose gels by electrophoresis for 15 hours at 1.5 V.cm^{-1} . DNA fragments were transferred to nitrocellulose filters by Southern blotting in X6 SSC (0.15M sodium chloride, 0.015M trisodium citrate is X1 SSC) and the filters hybridised to various probes labelled with ^{32}P by nick translation (7,8,9). Probes used included: JW 101 (α), JW 151 (γ), P ϵ 1.3 (ϵ) (10,11). Genomic Pst β and Pst δ segments in pBR322 were used for β and δ probes and the Bgl II 4.2 kb fragment electroeluted from λ HyG4 served as the $\psi\beta 1$ probe (12). Hybridisation was carried out for 18-24 hours at 42°C in 5% dextran sulphate and filters were washed for 1 hour at room temperature in X2 SSC and 1 hour at 65°C in 0.1 X SSC (13). Filters used in rehybridisation studies were not preincubated for rehybridisation but were washed for 1 hour at 65°C in 5 mM Tris-pH 7.8 and 0.1% sodium dodecyl sulphate. Bands were detected by autoradiography (14) and sized by comparison with bands obtained in Hind III, Bam HI digested λ DNA and normal DNA.

RESULTS

Haematological data and haemoglobin analysis Haematological data and haemoglobin analysis in the propositus showed a marginally low haemoglobin with a microcytic hypochromic blood

picture. The HbA₂ level was normal and the percentage of HbF raised to 8%, having previously been documented at 13% in mid-pregnancy. The ratio of G γ to A γ chains in HbF from SG was 0.75. Haematological data from a sister of SG (AT 21) showed a similar hypochromic microcytic anaemia. HbA₂ levels from all 5 members of SG's family were normal. The percentage of HbF was slightly raised in 4 out of 5 family members including the patient's son (AT 31) 2.3%, her two sisters (AT 20, 21) 1.1 and 2.0% and her husband (AT 32) 1.1%. The only member of the family with a percentage of HbF which fell within the normal range for the laboratory in Vanuatu was AT 19, a brother of SG (Table 1).

Restriction endonuclease analysis Normal band patterns were obtained in all but two instances when DNA from SG was digested with a wide variety of restriction enzymes and hybridised to γ , β , δ , ϵ and $\psi\beta 1$ probes. The only abnormal bands to be found were with the Bgl II and Bcl I restriction enzymes, which produced in addition to the normal 13 kb and 20 kb bands, larger fragments of 18 and 25 kb respectively, when hybridised to the γ probe (Table 2). These abnormal bands were not partial digestion products because they were consistently obtained on a number of occasions

Table 1 Haematological data and haemoglobin analyses for SG and family.

	HbA ₂ %	HbF % **
Propositus (SG)*	1.6	7.8
AT 21 (sister)	2.2	2.0
AT 20 (sister)	2.4	1.1
AT 19 (brother)	2.6	0.7
AT 31 (son)	1.8	2.3
AT 32 (husband)	2.2	1.1

* Hb 11.4 g/dl

MCV 71 fl, MCH 25.2 pg

Reticulocytes 131 x 10⁹/l

s ferritin 105 μ g/l

** normal range is

< 0.73% (mean + 2 S.D.).

Table 2 Size of globin gene fragments in restriction endonuclease digests of normal DNA and the propositus SG.

Enzyme	Normal DNA (kb)	DNA from SG (kb)
<u>Detected with γ probe</u>		
Bgl II	13	13, 18*
Bcl I	20	20, 25
Eco RI	7.2, 1.6, 2.7	7.2, 1.6, 2.7*
Bam HI	2.6, 5.1, 15.5	2.6, 5.1*, 15.5
Xba I	3.7, 5.1, 7.5	3.7, 5.1*, 7.5
Pst I	4.0, 5.1	4.0, 5.1*
Hind III	8.0, 3.5	8.0, 3.5*
Sac I	2.7, 3.8, 1.2	2.7, 3.8*, 1.2
Hpa I	5.1, 14	5.1, 14
Xho I	5.1	5.1
Pvu II	1.7, 4.6	1.7, 4.6*
<u>Detected with $\delta + \beta$ probe</u>		
Bgl II	7.7, 5.2	7.7, 5.2
Eco RI	2.3, 1.8, 5.2, 3.6	2.3, 1.8, 5.2, 3.6
Bam HI	15.5, 4.4, 1.8, 8.3	15.5, 4.4, 1.8, 8.3
Xba I	10.8	10.8
Pst I	2.3, 4.4	2.3, 4.4
Hind III	15, 7.5	15, 7.5
Hpa I	1.9, 1.4, 7.0	1.9, 1.4, 7.0
Pvu II	15	15
Taq I	2.9, 3.0, 8.5	2.9, 3.0, 8.5
<u>Detected with ϵ probe</u>		
Bgl II	10.5	10.5
Bcl I	12	12
Eco RI	4.0	4.0
Bam HI	13.3	13.3
Xba I	3.8	3.8
Pst I	14	14
Hind III	8	8
Hpa I	24	24
<u>Detected with $\psi\beta 1$ probe</u>		
Bgl II	4.2	4.2
Bcl I	7.5	7.5
Eco RI	7.2	7.2

* these bands demonstrated increased intensity.

with different blood samples. Furthermore, filters that were hybridised with γ and subsequently rehybridised with $\beta + \delta$ and ϵ probes failed to show evidence of partial digestion products in the Bgl II tracks. The 18 kb Bgl II and 25 kb Bcl I bands therefore represented an additional 5 kb of DNA between the Bgl II and Bcl I sites. The inserted DNA was most likely a third γ gene because various enzymes with restriction sites located within the Bgl II and Bcl I sites (Eco RI, Bam HI, Xba I, Pst I, Hind III, Sac I, Hpa I, Xho I and Pvu II) did not show any abnormal bands when hybridised to the γ probe. The absence of additional bands in the latter would only be possible if an exact copy of normal DNA had been inserted between the Bgl II sites so that all restriction enzyme sites retained their relative positions (Figure 1).

Further evidence for this arrangement came from the finding of differential intensity in bands hybridised with the γ probe. Duplication of normal band patterns was clearly seen by increased intensity of the 18 kb (Bgl II), 2.7 kb (Eco RI), 5.1 kb (Bam HI, Xba I, Pst I), 3.5 kb (Hind III), 3.8 kb (Sac I) and 4.6 kb (Pvu II) bands (Table 2, Figure 2).

Differential intensity was quantitated by comparing bands of equivalent size from the patient SG (S) and a normal control (N). The bands were standardised for the effect of DNA concentration and blotting by first relating their intensities to other bands within the same blots, for example, the 3.5 kb Hind III band in DNA from SG was more intense than the 8 kb band by a factor of 2.35. In normal DNA the same comparison between the 3.5 and 8 kb bands was 1.54, which thus gave an overall increased intensity of 1.5 for the 3.5 kb band in SG. Densitometry traces of the bands depicted in Figure 2 showed ratios between SG and normal of 1.1 (Bgl II), 1.6 (Pst I), 1.5 (Hind III) and 1.4 (Eco RI) (Figure 3). The results are close to the predicted 1.5 which would be found if there was an additional γ gene to which the γ probe could hybridise. The Bgl II bands gave a much lower ratio but this may be partly the effect of differential blotting between the 18 and 13 kb bands.

The likelihood that the 18 kb Bgl II and the 25 kb Bcl I bands were the result of a deletion small enough to have an

Figure 1. Normal β globin gene cluster and rearrangement for the triplicated γ gene. Bc Bcl I, Hd Hind III, E Eco RI, P Pst I, X Xba I, B Bam HI, Bg Bgl II, S Sac I, Pv Pvu II, Xh Xho I. Normal restriction sites were derived from references 15, 16, 17, 22 and our own observations for the 1.7 kb and 4.6 kb bands with Pvu II. The duplicated segment of DNA in the γ gene rearrangement (assuming the cross-over event occurred between 2 γ genes - see Discussion) is indicated .

undetectable effect on the 15.5 kb Bam HI, 14 kb Hpa I and the 7.5 kb Xba I bands, but deleting the Bgl II and Bcl I sites located in close proximity at the 3' end of the $A\gamma$ gene, was excluded by showing normal 4.2 kb Bgl II and 7.5 kb Bcl I fragments when these digests were hybridised to the $\psi B1$ probe.

Two possible gene arrangements would give rise to triplicated γ genes. These are: $G\gamma$ - $G\gamma$ - $A\gamma$ or $G\gamma$ - $A\gamma$ - $A\gamma$. There is only one difference in the restriction enzyme sites of the $G\gamma$ and $A\gamma$ genes and it is found at the 3' end of the third exon. In $G\gamma$ it is a Sac I site whilst in $A\gamma$ it is a Pst I site (Figure 1). The rearrangement can therefore be represented as: $G\gamma(\text{Sac})$ - $G\gamma(\text{Sac})$ - $A\gamma(\text{Pst})$ or $G\gamma(\text{Sac})$ - $A\gamma(\text{Pst})$ - $A\gamma(\text{Pst})$. Band

Figure 2. Bands obtained with various restriction endonuclease enzymes (Bgl II, Pst I, Hind III, Bam HI and Eco RI) and γ probe (upper) and $\beta + \delta$ probes (lower). S - patient SG, N - normal DNA.

patterns expected from the two alternatives would be: Pst I - 4, 5.1, 5.1, 0.9 kb or 4, 5.1, 0.9, 4.2, 0.9 kb. Sac I - 2.7, 3.8, 3.8, 1.2 kb or 2.7, 3.8, 1.2, 3.8, 1.2 kb. Since the 4.2

Figure 3. Densitometry traces of bands in Figure 2 showing increased intensity of the 3.5 kb Hind III, 18 kb Bgl II and 5.1 kb Pst I bands in patient SG. These traces were obtained by measuring at 590 nm the optical density of the various band patterns.

kb Pst I band was not detected (Table 2) the rearrangement must be G γ -G γ -A γ . Also consistent with this was the finding of greater intensity for the 5.1 kb (Pst I) and 3.8 kb (Sac I) bands. Furthermore, within the triplicated gene rearrangement the inserted DNA must be an exact duplication (in length and orientation) of existing sequences, otherwise abnormal bands would have been detected with the many restriction enzymes which have recognition sites in and around the γ genes. The ϵ , β and δ globin genes remain unaltered in the propositus SG as no abnormal bands were detected following mapping of these loci with a number of restriction enzymes (Table 2).

DNA samples from members of SG's family were also studied. Digestion with the enzyme Bgl II and hybridisation with the α probe produced bands of 12 and 7.4 kb in the normal DNA ($\alpha\alpha/\alpha\alpha$). SG and her three sibs demonstrated only a 15.7 kb band consistent with the $-\alpha^{3.7}/$ haplotype. Since the α thalassaemia-1 haplotype ($--/$) has not been found in this population (D.K.Bowden et al, in preparation) the most likely genotype for the above is

- α /- α . SG's 3 year old son (AT 31) showed bands at 15.7, 12 and 7.4 kb ($\alpha\alpha$ /- α) while her husband (AT 32) was normal ($\alpha\alpha$ / $\alpha\alpha$). Screening the family with Bgl II and hybridising to the γ probe showed that 2 out of 3 sibs, the patient's son and her husband were heterozygous for the 3 γ gene haplotype ($\gamma\gamma\gamma$ / $\gamma\gamma$). The one member of the family with a normal $\gamma\gamma$ / $\gamma\gamma$ genotype was the brother (AT 19) who was also the only individual to have a normal percentage of HbF. DNA from 27 unrelated individuals from Aoba (the island from where SG's husband originated) were also screened and 2 were positive for the $\gamma\gamma\gamma$ / haplotype.

DISCUSSION

Restriction mapping studies have demonstrated the genotype $G\gamma G\gamma A\gamma$ / $G\gamma A\gamma$ in 5 out of 6 living members of a native family and 2 out of 27 unrelated individuals from Vanuatu. The rearrangement $G\gamma$ - $G\gamma$ - $A\gamma$ represents the first example in which the γ gene complex has been altered by an insertion of additional albeit normal DNA, rather than by a deletion.

The mechanism by which the $\gamma\gamma\gamma$ / haplotype might arise is of some interest. The various gene rearrangements found in the α^+ thalassaemia (α thalassaemia-2) deletion haplotypes (- α /) are thought to have resulted from interchromosomal crossing-over between areas of homology (18,19). The finding of chromosomes that contain the corresponding triplicated α gene ($\alpha\alpha\alpha$ /) haplotypes lends some support to this theory (20, 21). Recent sequence analysis of the $G\gamma$ and $A\gamma$ genes has shown highly conserved areas of DNA sequence in the 5' flanking, 5' non-coding, the complete coding sequence, all of IVS 1 and most of the larger IVS 2 (22). Slightom et al (22) have demonstrated a region of 'simple sequence' DNA 3' to an invariant region in the centre of IVS 2. This 'hot spot' provides an area through which DNA exchange may occur and the authors have ascribed to this 'hot spot' a possible role in gene conversion. Alternatively such an area would provide a site for unequal crossing-over between homologous chromosomes and if the $G\gamma$ and $A\gamma$ genes were to be involved in this cross-over the resultant hybrid chromosome would be $G\gamma$ - $G\gamma$ - $A\gamma$.

The coexistence of different α thalassaemia genotypes in SG

and her family precludes a definitive characterisation of the haematological phenotype associated with the duplicated G γ locus. Thus the microcytic hypochromic anaemia in SG and sister would be consistent with homozygous α^+ thalassaemia ($-\alpha/-\alpha$) and the additional γ gene may contribute little if anything to the anaemia. It is also likely that the raised HbF of 8 % in the propositus SG is not directly attributable to the $\gamma\gamma\gamma/$ haplotype per se, as other individuals with the same γ globin gene rearrangement fail to show similar elevations in HbF. The presence of a heterocellular HPFH gene which has been reported in low frequency in many populations may be an alternative explanation for the 8% HbF level in SG (4). On the other hand, the marginal elevation of HbF found in 4 members of this family may be related to the $\gamma\gamma\gamma/$ rearrangement, although this will require additional numbers for confirmation as the levels of HbF in 2 cases were only slightly above the normal range. Furthermore subjects without the α thalassaemia gene will be needed to resolve the above, as a possible effect of α thalassaemia on the HbF level cannot be excluded.

Although a number of examples of the $\gamma\gamma\gamma/$ haplotype have been detected in Vanuatu, the significance of this rearrangement remains unclear. To date a limited survey in the Sardinian and Saudi Arab populations has found only 2 further examples of the $\gamma\gamma\gamma/$ haplotype and these were present in the same Sardinian family. Until recently the inhabitants of the various islands which make up Vanuatu have remained isolated. Local inbreeding in such communities could therefore explain the relatively high incidence of the $\gamma\gamma\gamma/$ haplotype in this population. Whether there are other factors favouring selection of this gene rearrangement remains to be determined.

ACKNOWLEDGEMENTS

Pst β and δ subclones were gifts from Dr. T. Maniatis. We thank Dr. R. Cobcroft, Princess Alexandra Hospital, Qld. Australia, for his assistance and Dr. A. Cao, University of Cagliari, Italy for providing blood samples from Sardinian patients. This work was supported by a grant from the Rockefeller Foundation.

REFERENCES

1. Flavell, R.A., Kooter, J.M., De Boer, E., Little, P.F.R. and Williamson, R. (1978). *Cell* **15**, 25-41.
2. Little, P.F.R., Flavell, R.A., Kooter, J.M., Annison, G. and Williamson, R. (1979). *Nature* **278**, 227-231.
3. Bernards, R., Little, P.F.R., Annison, G., Williamson, R. and Flavell, R.A. (1979). *Proc. Natl. Acad. Sci. USA* **76**, 4827-4831.
4. Weatherall, D.J. and Clegg, J.B. (1981). *The Thalassemia Syndromes*. 3rd ed. Blackwell Scientific Publications, Oxford.
5. Clegg, J.B., Metaxatou-Mavromati, A., Kattamis, C., Sofroniadou, K., Wood, W.G. and Weatherall, D.J. (1979). *Brit. J. Haemat.* **43**, 521-536.
6. Gross-Bellard, M., Oudet, P. and Chambon, P. (1973). *Eur. J. Biochem.* **36**, 21-38.
7. Southern, E.M. (1975). *J. Mol. Biol.* **98**, 503-517.
8. Kan, Y.W. and Dozy, A.M. (1978). *Proc. Natl. Acad. Sci. USA* **75**, 5631-5635.
9. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975). *Proc. Natl. Acad. Sci. USA* **72**, 1184-1188.
10. Wilson, J.T., Wilson, L.B., de Riel, J.K., Villa-Komaroff, L., Efstratiadis, A., Forget, B. and Weissman, S.M. (1978). *Nucl. Acid Res.* **5**, 563-581.
11. Baralle, F., Shoulders, C.C., Goodbourn, S., Jeffreys, A. and Proudfoot, N.J. (1980). *Nucl. Acid Res.* **8**, 4393-4404.
12. Fritsch, E.F., Lawn R.M. and Maniatis, T. (1980). *Cell* **19**, 959-972.
13. Orkin, S.J., Old, J., Lazarus, H., Altay, C., Gurgey, A., Weatherall, D.J. and Nathan, D.G. (1979). *Cell* **17**, 33-42.
14. Laskey, R.A. and Mills, A.D. (1977). *FEBS Letters* **82**, 314-316.
15. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1979). *Nature* **279**, 598-603.
16. Tuan, D., Biro, P.A., de Riel, J.K., Lazarus, H. and Forget, B.G. (1979). *Nucl. Acid Res.* **6**, 2519-2544.
17. Van der Ploeg, L.H.T., Konings, A., Oort, M., Roos, D., Bernini, L. and Flavell, R.A. (1980). *Nature* **283**, 637-642.
18. Lauer, J., Shen, C-K, J. and Maniatis, T. (1980). *Cell* **20**, 119-130.
19. Embury, S.H., Miller, J.A., Dozy, A.M., Kan, Y.W., Chan, V. and Todd, D. (1980). *J. Clin. Invest.* **66**, 1319-1325.
20. Higgs, D.R., Old, J.M., Pressley, L., Clegg, J.B. and Weatherall, D.J. (1980). *Nature* **284**, 632-635.
21. Trent, R.J., Higgs, D.R., Clegg, J.B. and Weatherall, D.J. (1981). *Brit. J. Haemat.* **49**, 149-152.
22. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980). *Cell* **21**, 627-638.