Proc. Natl. Acad. Sci. USA
Vol. 75, No. 12, pp. 5950–5954, December 1978
Biochemistry

The duplicated human α globin genes lie close together in cellular DNA

(restriction enzyme mapping/molecular hybridization/gene linkage/α-thalassemia)

STUART H. ORKIN*

Division of Hematology and Oncology, Children's Hospital Medical Center, and the Sidney Farber Cancer Center, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115

Communicated by David Baltimore, September 25, 1978

ABSTRACT **The organization of α globin genes in normal human DNA was examined by restriction endonuclease mapping. α globin-specific fragments in endonuclease digests of total cell DNA were identified after electrophoresis by hybridization with [^{32}P]cDNA following the blotting procedure of Southern [(1975) *J. Mol. Biol.* 98, 503–517]. The data provide direct evidence for the duplication of α genes and further indicate that these loci are closely linked within a single restriction fragment. The *Hin*dIII sites (codons 90/91) of these duplicated genes lie approximately 3.7 kilobases apart in the physical map proposed for this region. This organization of α genes can be altered in DNA of individuals with α-thalassemia.**

The molecular genetics of human hemoglobins has been the subject of intensive study (1, 2). All normal hemoglobins are formed as tetramers of two α and two non-α (γ, δ, β) globin chains. Genetic studies of variant hemoglobins have provided information regarding the number and linkage relationships of the structural genes for these globins. The β and δ loci are present as single copies per haploid genetic complement and are genetically linked to one another and to duplicated γ loci on a single chromosome recently proposed to be number 11 (3). Most evidence favors duplication of α loci, which are unlinked to the non-α genes, and assigned to chromosome 16 (4–7).

To define more precisely the number and arrangement of α globin genes in cell DNA, I have performed a series of restriction endonuclease mapping experiments. In this work α-specific fragments in restriction enzyme digests of total DNA have been identified after electrophoresis by hybridization with [^{32}P]cDNA probes following the blotting procedure introduced by Southern (8). The data indicate that the α genes are normally duplicated and lie quite close together in a single region of the DNA. This arrangement of α genes can be disturbed in pathologic states of α globin expression, known as α-thalassemias.

MATERIALS AND METHODS

Globin RNAs and Cell DNAs. Human globin mRNA was isolated as previously described (9). α globin mRNA, about 90% free of β sequences, was prepared by electrophoresis in formamide (10). High molecular weight cell DNA was prepared from cultured lymphocyte cell lines of normal donors (11).

Globin cDNA Probes. Globin cDNAs were synthesized at a specific activity of 4×10^8 cpm/μg by using [^{32}P]dGTP as the labeled nucleotide (9). α cDNAs contained approximately 50% full-length transcripts and smaller, heterogeneous material when examined in 3% agarose slab gels. Short cDNA, enriched for 3′-specific mRNA sequences, was prepared by using suboptimal synthesis conditions, including 1.5 μM dGTP and limiting avian myeloblastosis virus DNA polymerase. Material less then 300 nucleotides in length was utilized as 3′-specific probe. In the purification of short, unlabeled cDNA, tracer short mouse globin cDNA was added as an internal size marker and material less than 300 nucleotides was isolated.

Restriction Enzymes and Agarose Electrophoresis. Enzymes were purchased from either New England BioLabs or Boehringer Mannheim. Digestions with *Eco*RI were performed in 100 mM Tris·HCl, pH 7.5/50 mM NaCl/10 mM $MgCl_2$/1 mM dithiothreitol to retain specificity for cleavage of G-A-A-T-T-C (12). Buffers recommended by the commercial suppliers were used in other digestions. DNA samples were phenol extracted and ethanol precipitated prior to electrophoresis in agarose (13).

Identification of α Globin-Specific DNA Fragments. After electrophoresis DNA fragments were transferred to Millipore filter sheets by blotting (8, 14). Hybridization on filters was carried out as described by Botchan *et al.* (15) with 2–3 ng of [^{32}P]cDNA probe per ml of hybridization solution for 24 hr at 68°C. In most instances α cDNA was employed as the probe to detect α-specific DNA fragments because minor contaminant β cDNA in α cDNA does not hybridize appreciably under these conditions. In some experiments (see Fig. 1) mixed α and β cDNA was used as the probe after purification of *Eco*RI-digested cell DNA free of non-α globin sequences. After hybridization, filters were washed under stringent conditions, including at least two 1-hr rinses in 15 mM NaCl/1.5 mM sodium citrate at 65°C. Autoradiography of dried filters was performed with calcium tungstate intensifying screens.

RESULTS

The specificity of several restriction enzymes for cleavage within known α gene sequences has been determined by analysis of their activity on *in vitro* synthesized, double-stranded DNA prepared from mRNA (9) and by sequencing cloned cDNA (16). *Eco*RI, *Bam*HI, *Hpa* I, and *Sal* I do not cleave α sequences, whereas *Hin*dIII and *Hin*dII introduce single cuts at codons 90/91 and 96/97, respectively.

α-Specific fragments in restriction endonuclease digests of total cell DNA were identified by molecular hybridization after electrophoresis in agarose and blotting onto filter sheets. A summary of the observed fragments is given in Table 1. The physical map of the DNA region containing α sequences deduced from these data is shown in Fig. 1. The data and considerations leading to this model are presented below.

Initial Evidence for α Gene Duplication and Linkage. In *Eco*RI-digested DNA, α sequences are normally confined to a single fragment of approximately 14 megadaltons (ref. 13 and Fig. 2). If more than one copy of the α sequence per haploid

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. §1734 solely to indicate this fact.

* Address correspondence to S.H.O. at the Division of Hematology and Oncology, Children's Hospital Medical Center, 300 Longwood Avenue, Boston, MA 02115.

Table 1. α-Specific DNA fragments, megadaltons

	–	+ *Sal* I	+ *Bam*HI	+ *Hin*dIII	+ *Bam*HI + *Hin*dIII	+ *Hin*dIII + *Hpa* I	+ *Hin*dIII + *Hpa* I	+ *Hin*dII
*Eco*RI	≈14	≈14	8–8.5	≈10 (5′)	4.5	3.0	2.1	2.1
				2.3	2.3	2.5	1.85	1.85
				(5′ and 3′)	1.3		1.3	0.9
				1.3 (3′)			0.5	0.5
*Hin*dIII	≈10		4.5			2.8		
	2.8		2.3			2.1		
	2.3		1.9			1.85		
						0.5		

The sizes of α-specific DNA fragments were determined by reference to fragments of λ phage DNA digested with *Eco*RI or *Hin*dIII and simian virus 40 DNA treated with *Hae* III or *Hin*dIII. The orientation of the *Eco*RI + *Hin*dIII fragments was determined by use of short, 3′-specific cDNA as probe or competitor (see Fig. 6).

genome existed, we would expect to observe more than one α-specific fragment after treatment with other enzymes that do not cut α sequences, unless *Eco*RI and other restriction sites were located in a repeated fashion about α loci or the DNA sequences between two α genes did not contain sequences recognized by the enzymes tested. Initial evidence for α gene duplication was provided by digestion with *Hpa* I. When *Eco*RI-digested DNA was treated with *Hpa* I, two fragments, about 3.0 and 2.5 megadaltons, were present (Fig. 2). In contrast, *Bam*HI digestion produced one fragment containing α sequences of about 8–8.5 megadaltons (Fig. 2). In other experiments (not shown), *Sal* I did not appear to cleave the large *Eco*RI α-specific fragment, whereas *Sst* I, another enzyme that does not cut α sequences (unpublished data), produced two fragments (2.9 and 2.4 megadaltons).

Because *Hin*dIII cuts α sequences at a site centrally placed in the nucleotide sequences represented in the non-poly(A) portion of mature α mRNA (9), DNA digested with this enzyme would be expected to contain two α-specific fragments if there were only a single α gene per haploid complement, three fragments if there were closely linked, duplicated loci, and four fragments if duplicated loci were widely separated or separated by a *Hin*dIII site outside an α sequence. In accord with the predictions of a two-gene model in which close physical linkage is present, *Hin*dIII digestion yielded three α-specific fragments, about 10, 2.8, and 2.3 megadaltons (Fig. 3).

Ordering of Restriction Sites in the Physical Map. To position restriction sites about the α loci it is useful to consider the three *Hin*dIII fragments individually.

The 2.8-megadalton fragment was cleaved by *Bam*HI (Fig. 3) and *Eco*RI (Fig. 4) but not *Hpa* I (Fig. 5). Because this fragment is cleaved by *Eco*RI, it traverses one end of the 14-megadalton *Eco*RI α-specific fragment and is drawn to the left in the model illustrated in Fig. 1. The *Bam*HI site in this *Hin*dIII fragment lies to the left of this end *Eco*RI site on the basis of comparison of the *Hin*dIII + *Bam*HI and *Hin*dIII + *Eco*RI digests (see Table 1).

The 10-megadalton *Hin*dIII fragment is cleaved by *Bam*HI (Fig. 3) but not *Eco*RI (Fig. 4) and, therefore, overlaps the *Bam*HI site present in the large *Eco*RI fragment. This large *Hin*dIII fragment is depicted on the right of the physical map in Fig. 1. It should be noted that this *Hin*dIII fragment is not well seen in some experiments for two reasons. First, by virtue of its large size, it is inefficiently transferred during the blotting procedure (8). Second, it appears to carry 5′ α coding sequences (see below), which are underrepresented in cDNA preparations synthesized with oligo(dT) as primer.

The third *Hin*dIII fragment, about 2.3 megadaltons, lies between the other two because it is cleaved by neither *Eco*RI nor *Bam*HI (Figs. 3 and 4). However, it is cut by *Hpa* I, which I propose cleaves between the duplicated α loci (Fig. 5). No additional *Hin*dIII sites appear to be present between the *Eco*RI site on the left and the *Bam*HI site depicted to the right in the physical map, because the entire stretch of DNA between these positions (8–8.5 megadaltons) can be accounted for in *Eco*RI + *Hin*dIII + *Bam*HI digests (see Table 1).

The results of digestion with *Hin*dII + *Eco*RI (see Table 1), in which three of the observed fragments comigrated with the

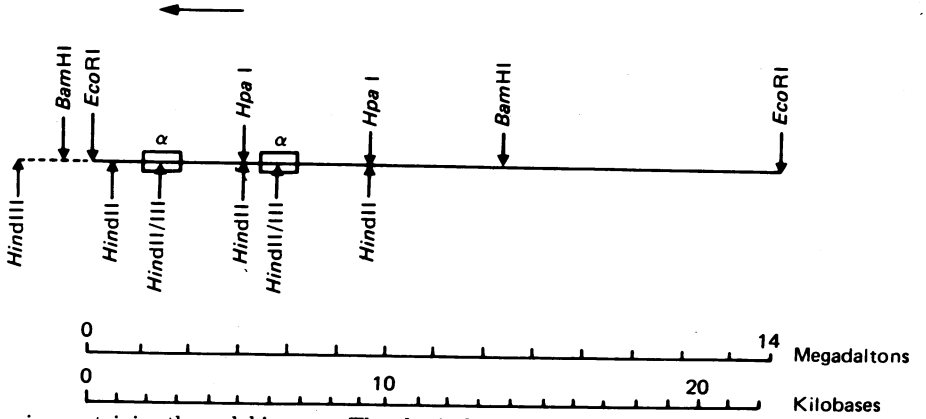

FIG. 1. The DNA region containing the α globin genes. The physical map was constructed using the sizes of the fragments summarized in Table 1. Restriction sites are positioned with respect to the *Eco*RI site within the 2.8-megadalton *Hin*dIII fragment. The positions of the codons 90–96 of the α loci are defined by *Hin*dIII and *Hin*dII sites, which are shown together due to their proximity on the indicated scale. The actual size of each α gene is arbitrarily depicted by the boxes. The orientation of the map is shown by the arrow drawn from 5′ to 3′ of the α sequence (as represented in mRNA).

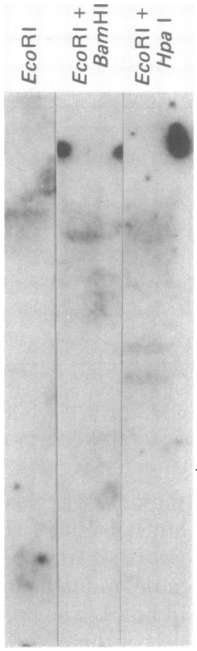

FIG. 2. α-Specific DNA fragments in *Eco*RI-digested DNA treated with *Bam*HI or *Hpa* I. *Eco*RI-digested normal DNA of 10–15 megadaltons was initially prepared by electrophoresis and was re-treated with *Eco*RI, *Bam*HI, or *Hpa* I prior to electrophoresis in 0.85% agarose. Approximately 0.75–1 μg of DNA was applied to each gel lane. α-Specific fragments were identified by hybridization with mixed α and β cDNA. Only α-specific fragments are evident because all β-specific *Eco*RI fragments are less than 5 megadaltons (13) and were removed in the preparation of 10- to 15-megadalton *Eco*RI DNA. The single bands visible in the *Eco*RI and *Bam*HI lanes are 14 and 8–8.5 megadaltons, respectively. The two *Hpa* I bands are 3.0 and 2.5 megadaltons. The smudge at about 8–9 megadaltons in the *Hpa* I lane is nonspecific contamination (see Fig. 4).

"internal" *Eco*RI + *Hin*dIII + *Hpa* I fragments, provide confirmation of the locations of *Hin*dIII sites within two α sequences and their positions relative to the *Hpa* I sites. This is the case because the *Hin*dII cleavage sequence G-T-pyrimidine-purine-A-C is recognized by *Hpa* I (17) and is also located within the α gene sequence in close proximity to the *Hin*dIII site (codons 96/97 versus 90/91).

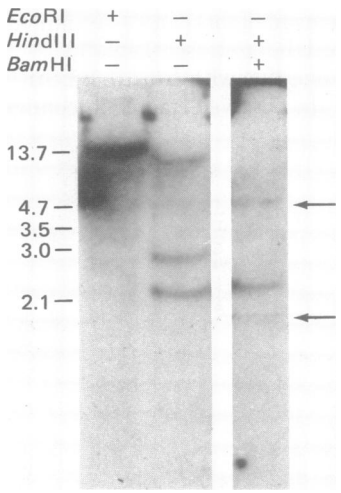

FIG. 3. α-Specific fragments in digests with *Hin*dIII with and without *Bam*HI. Fifteen micrograms of DNA was applied to the *Eco*RI and *Hin*dIII lanes, and 10 μg to the *Hin*dIII + *Bam*HI lane. Electrophoresis was in 1.4% agarose and fragments were detected with α cDNA. Markers in megadaltons are given to the left. Arrows indicate the 4.5- and 1.9-megadalton bands in the *Hin*dIII + *Bam*HI digest.

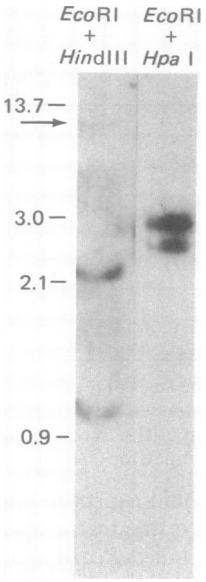

FIG. 4. Presence of an *Eco*RI site within the 2.8-megadalton *Hin*dIII fragment. Approximately 15 μg of each DNA sample was electrophoresed in 1.4% agarose. α cDNA was used as the probe. The *Eco*RI + *Hpa* I lane from the same gel is included for reference. The arrow indicates the largest *Hin*dIII fragment in the *Hin*dIII + *Eco*RI digest.

Further Support for Gene Duplication and the Orientation of the α Loci. Although the results of *Hpa* I digestion suggest the presence of duplicated α loci, they are not sufficient alone to exclude a one-gene model in which a cellular α gene contains at least two large intervening sequence regions (18), one of which has an *Hpa* I site and the other a *Hin*dIII site. However, experiments using short 3′-specific cDNA either as label or

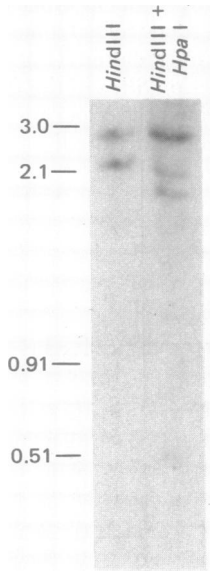

FIG. 5. α-Specific fragments in *Hin*dIII + *Hpa* I-digested DNA. Fifteen micrograms of DNA was applied to each lane of a 1.4% gel and hybridized with α cDNA. The positions of the two smallest *Hin*dIII fragments are included for reference in addition to the molecular weight markers on the left.

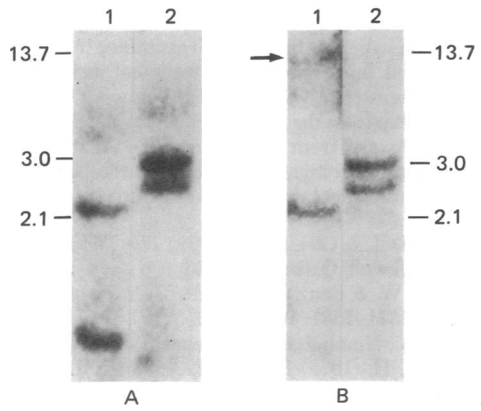

FIG. 6. Use of short, 3′-enriched cDNA as probe or competitor. Lanes 1, normal DNA digested with *Eco*RI + *Hin*dIII; lanes 2, normal DNA digested with *Eco*RI + *Hpa* I. Electrophoresis was in 1.4% agarose. (*A*) Hybridization with short, 3′-specific α cDNA. Note the absence of the largest *Eco*RI + *Hin*dIII fragment in lane 1. (*B*) Hybridization with α [^{32}P]cDNA mixed with a 10-fold excess of nonradioactive, short (3′-specific) α cDNA. Note the absence of the band at about 1.3 megadaltons. The arrow points to the position of the largest *Hin*dIII + *Eco*RI fragment, which is barely visible in the photograph but discernible above background nonspecific radioactivity in the region below in the original autoradiogram. Note especially overall the nearly equal labeling of the two *Eco*RI + *Hpa* I bands in both *A* and *B*, and, in marked contrast, the unequal labeling of the 1.3-megadalton fragment.

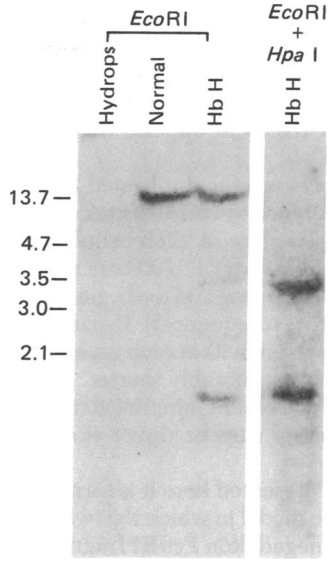

FIG. 7. Abnormal α-specific fragments in Hb H DNA. On the left, hybridization of normal *Eco*RI-digested DNA is compared with hybridizations with homozygous α-thalassemia (hydrops fetalis) DNA and Hb H DNA obtained from the lymphocyte line (kindly provided by H. Lazarus) of a Mediterranean patient. Note the complete deletion of α sequences in the hydrops fetalis sample, as previously described (13). The molecular mass of the smaller *Eco*RI fragment in the Hb H DNA is about 1.6 megadaltons. Faint bands at 3.5, 2.3, and 1.3 megadaltons, seen more prominently in the Hb H DNA lane than in the normal or hydrops lanes, are β-specific fragments detected by residual contamination of α cDNA with β cDNA in this experiment. The molecular masses of the two *Eco*RI + *Hpa* I bands present in this Hb H sample, shown on the right, are 3.5 and 1.6 megadaltons.

competitor (Fig. 6) provide strong arguments against this possibility. When 3′-specific cDNA of less than 300 nucleotides in length was used as probe, the two α-specific *Hpa* I + *Eco*RI bands were approximately equally labeled under conditions in which only the two smallest *Eco*RI + *Hin*dIII bands were seen (Fig. 6*A*). When nonradioactive 3′-specific cDNA was used to block hybridization of 3′-specific sequences in the presence of labeled α cDNA, the *Hpa* I bands were again equally labeled even though the smallest *Eco*RI + *Hin*dIII fragment (1.3 megadaltons, derived from the left *Hin*dIII fragment in the map) was not visualized (Fig. 6*B*). The lack of differential hybridization of the *Hpa* I α-specific bands with these probes indicates that they contain similar α gene contents and do not arise from cleavage of a single α gene within an intervening sequence recognized by *Hpa* I. If the latter were the case, the *Hpa* I fragments would contain α sequences from different halves of the gene sequence and would have been differentially labeled by these probes.

On the basis of these results, 3′ coding sequences have been assigned to the smallest *Eco*RI + *Hin*dIII fragment (1.3 megadaltons, derived from the 2.3-megadalton *Hin*dIII fragment), both 3′ and 5′ coding sequences to the middle (2.3-megadalton) fragment, and only 5′ sequences to the largest *Hin*dIII fragment (10 megadaltons). Because of the difficulties in preparing entirely 5′-specific probes from cDNA and in achieving efficient transfer of large molecular weight DNA, it has not been possible to demonstrate more conclusively the 5′ coding sequence within this latter fragment. The duplicated α globin genes lie in the same orientation in the DNA approximately 2.3 megadaltons (3.7 kilobases) apart as measured from *Hin*dIII sites located at codons 90/91 of the coding sequence.

Abnormal α Gene Organization in α-Thalassemia. Studies in Asians have suggested that deletion of α genes is the predominant molecular defect in α-thalassemias, conditions in which α globin synthesis is deficient (2, 6, 19–21). Hemoglobin H (Hb H) disease, an intermediate form of α-thalassemia, is currently thought to be associated almost always with deletion of three of the four normal α genes per diploid cell (6). We have recently examined several Hb H DNAs after digestion with *Eco*RI and *Eco*RI + *Hpa* I. In nearly all cases, as expected, an abnormal pattern of α-specific fragments has been observed (unpublished data). However, we have already encountered many instances in which more than one α gene is present. An example of one such result is shown in Fig. 7. In this particular Hb H DNA, two α-specific *Eco*RI fragments were visible, (about 14 and 1.6 megadaltons), and two *Eco*RI + *Hpa* I fragments (3.5 and 1.6 megadaltons), neither of which was present in normal DNA. Although in classical instances of Hb H disease we would expect to observe only a single *Eco*RI or *Eco*RI + *Hpa* I fragment, indicative of one gene copy per cell, this Hb H disease patient has at least two α genes. Hybridization kinetic analysis of this DNA with α cDNA provides independent support for the presence of two α loci (unpublished data).

DISCUSSION

The approach pioneered by Southern (8) to study specific sequences in complex DNAs has been applied here to the analysis of the human α globin genes. Studies on the distribution of hemoglobins containing mutant α chains within family members (4), the clinical features of α-thalassemias (5), and hybridization kinetic analysis in solution (6) have previously suggested that the α genes are normally duplicated. The data presented here provide direct physical evidence for α gene duplication and further demonstrate that the duplicated loci are closely linked within a small region of the DNA (Fig. 1). If the entire unit depicted in this map were repeated in the genome, any multiple of two α genes per haploid genetic complement might exist. In view of the genetic evidence, supported by solution hybridization assays, this appears unlikely to be the

case. The mapping data do argue strongly against the presence of additional, silent copies of the α loci. Because the principal features of this physical map of the α gene region have been observed in five normal and four β-thalassemic DNAs isolated from a variety of tissues, it appears that this organization is highly conserved.

The distance between the duplicated α genes has been defined here by reference to *Hin*dIII sites within the coding sequences. The precise size of each cellular α gene is not yet known. At a maximum about 3100 base pairs could reside between the end of one α gene and the beginning of the next [3700 base pairs $-$ ($2 \times \frac{1}{2}$ α sequence)]. If sizeable intervening sequences (18) are present within each gene, the actual intergene distance would be considerably shorter. A search for such sequences within α loci will require fine structure mapping using additional restriction sites or direct examination of cloned cellular genes.

With the data presented here it is formally possible to construct an alternate model in which the two α genes are situated on different 14-megadalton *Eco*RI fragments in the positions of the loci depicted in Fig. 1. This is considered extremely unlikely, however, because this would require identical positioning of *Bam*HI, *Hin*dIII, *Hpa* I, and *Hin*dII sites relative to the *Eco*RI sites in both fragments. The physical map proposed here suggests that the duplicated α loci should be so very tightly linked genetically that no recombination would occur in appropriate mating pairs. Unfortunately, insufficient genetic data are available to test this important prediction. In one suitable family where evidence suggested that duplicated α loci were not linked, the paternity of the offspring was in doubt (22). Using modifications of the methods employed here, others have examined the δ–β gene complex in human DNA and provided evidence for the physical linkage of these loci as well (23, 24).

Restriction mapping of DNA from α-thalassemic individuals will help define further the molecular defects in these disorders. Gene deletion is thought to be the primary defect in α-thalassemias (2, 6). However, in one form, Hb H disease, which is the consequence of α globin synthesis equivalent to that produced by a single normal α gene, deletion of three of the four normal α loci may not always occur. Indeed, poorly functional α loci, perhaps analogous to defective β loci in β-thalassemias, may be present at a higher frequency than previously suspected. The example of DNA from an individual with Hb H disease shown here in which at least two α genes are present (Fig. 7) provides direct evidence for the existence of a nondeletion α-thalassemic defect, previously proposed in a single Asian family on the basis of hybridization analysis in solution (21). Our results with additional Hb H DNA samples extend this finding and demonstrate more molecular heterogeneity in this form of α-thalassemia than previously suspected. Interaction of deletion and nondeletion defects may contribute to the clinical heterogeneity of α-thalassemias and the absence of the homozygous deletion state (hydrops fetalis) in non-Asian populations.

I am indebted to Dr. David G. Nathan for his continued encouragement, criticism, and support and to Dr. David Baltimore for his excellent comments and suggestions on this work. I especially wish to thank Dr. Jerry Schwaber for his unselfish assistance in the growth of the lymphocyte cell lines. The skilled technical assistance of Ms. Domenica Paci is most greatly appreciated. This work was supported in part by a fellowship from the Medical Foundation, a Basil O'Connor Starter Research Award from the National Foundation–March of Dimes, and a Young Investigator Award from the National Heart, Lung, and Blood Institute of the National Institutes of Health.

1. Weatherall, D. J. & Clegg, J. B. (1972) *The Thalassaemia Syndromes* (Blackwell, Oxford), 2nd Ed.
2. Nienhuis, A. W. & Benz, E. J., Jr. (1977) *N. Engl. J. Med.* **297**, 1318–1328; 1371–1381; 1430–1436.
3. Deisseroth, A., Nienhuis, A., Lawrence, J., Giles, R., Turner, P. & Ruddle, F. H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1456–1460.
4. Hollan, S. R., Szlenyi, J. G., Brimhall, B., Duerst, M., Jones, R. T., Koler, R. D. & Stocklen, Z. (1972) *Nature (London)* **235**, 47–50.
5. Lehnmann, H. (1970) *Lancet* ii, 78–80.
6. Kan, Y. W., Dozy, A. M., Varmus, H. E., Taylor, J. M., Holland, J. P., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1975) *Nature (London)* **255**, 255–256.
7. Deisseroth, A., Nienhuis, A., Turner, P., Velez, R., Anderson, W. F., Ruddle, F. H., Lawrence, J., Creagan, R. & Kucherlapati, R. (1977) *Cell* **12**, 205–218.
8. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
9. Orkin, S. H. (1978) *J. Biol. Chem.* **253**, 12–15.
10. Orkin, S. H., Swan, D. & Leder, P. (1975) *J. Biol. Chem.* **250**, 8753–8760.
11. Gross-Bellard, M., Oudet, P. & Chambon, P. (1973) *Eur. J. Biochem.* **36**, 21–38.
12. Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M. & Boyer, H. W. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3310–3314.
13. Orkin, S. H., Alter, B. P., Altay, C., Mahoney, J., Lazarus, H., Hobbins, J. C. & Nathan, D. G. (1978) *N. Engl. J. Med.* **299**, 166–172.
14. Cory, S. & Adams, J. M. (1977) *Cell* **11**, 795–805.
15. Botchan, M., Topp, W. & Sambrook, J. (1976) *Cell* **9**, 269–287.
16. Little, P., Curtis, P., Coutelle, C., Van Den Berg, J., Dalgleish, R., Malcolm, S., Courtney, M., Westaway, D. & Williamson, R. (1978) *Nature (London)* **273**, 640–643.
17. Nathans, D. & Smith, H. O. (1975) *Annu. Rev. Biochem.* **44**, 273–293.
18. Tilghman, S. M., Tiemeier, D. C., Seidman, J. C., Peterlin, B. M., Sullivan, M., Maizel, J. V. & Leder, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 725–729.
19. Ottolenghi, S., Lanyon, W. G., Paul, J., Williamson, R., Weatherall, D. J., Clegg, J. B., Pritchard, J., Pootrakul, S. & Boon, W. H. (1974) *Nature (London)* **251**, 389–392.
20. Taylor, J. M., Dozy, A. M., Kan, Y. W., Varmus, H. E., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1974) *Nature (London)* **251**, 392–393.
21. Kan, Y. W., Dozy, A. M., Trecartin, R. & Todd, D. (1977) *N. Engl. J. Med.* **297**, 1081–1084.
22. Rucknagel, D. L. & Winter, W. P. (1974) *Ann. N. Y. Acad. Sci.* **241**, 80–92.
23. Mears, J. G., Ramirez, R., Leibowitz, D. & Bank, A. (1978) *Cell* **15**, 15–23.
24. Flavell, R. A., Kooter, J. M., DeBoer, E., Little, P. F. R. & Williamson, R. (1978) *Cell* **15**, 25–41.