A common mutant *EcoRI* restriction endonuclease site in the 5' flanking portion of the human α -globin gene

(restriction mapping/thalassemia)

E. BEUTLER*, W. KUHL*, AND C. JOHNSON†

*Department of Clinical Research, Scripps Clinic and Research Foundation, La Jolla, California 92037; and †University of Southern California, Los Angeles, California 90033

Contributed by Ernest Beutler, July 27, 1981

ABSTRACT A mutant *EcoRI* endonuclease restriction site has been identified in 3 of 37 Black subjects and in 2 sibs of one of these persons. This mutation was not encountered in 13 Whites. It is located approximately 6 kilobases "inside" the normal site in the 5' flanking sequence of the α -globin chain complex. The shortened α -globin gene-bearing segment produced in the *EcoRI* digest produces a restriction map similar to that observed for the common α -globin gene deletion observed in the Black population. However, the restriction map with *Bam*HI is normal, confirming that all four α loci are present.

α -Thalassemia is a disorder in which the rate of synthesis of the α -globin chains of hemoglobin is impaired. In its most severe form, when none of the four normally present α loci are functioning, it invariably produces fetal or early neonatal death. In the milder forms, in which the functioning of only one or two α loci is impaired, there are no clinical manifestations and little or no anemia or microcytosis of the erythrocytes. The diagnosis can be established, however, by measuring the rate of synthesis of α - and β -globin chains in reticulocytes. An alternative method for the detection of the α -thalassemias has also recently become possible; with the availability of DNA probes for the globin genes, deletions can be detected by measuring the size of the globin gene-bearing fragments after digestion of the DNA with an appropriate restriction endonuclease.

Fig. 1 shows a map of the α -globin region (1). The restriction endonuclease *EcoRI* cleaves DNA at the palindrome 5'-G-A-A-T-T-C-3'. The *EcoRI* sites are outside the α -globin complex. When one of the α -globin genes has been deleted, the gene-bearing segment of DNA has a length of 19 kilobases (kb) instead of the normal 23 kb. This fact, and the relatively low cost of the *EcoRI* restriction endonuclease, has made this enzyme a convenient way of detecting the single gene deletion. This approach has been particularly useful in the diagnosis of α -thalassemia in the Black population. Here, a single α locus deletion has been found to be extremely common, with an estimated gene frequency of 0.16 (2). Thus, approximately 27% of the Black population would be expected to have the mild form of α -thalassemia in which a single α locus has been deleted.

We have recently determined the prevalence of α -thalassemia in the U.S. Black population by measuring globin synthetic ratios. A fair degree of concordance between the *EcoRI* restriction map and the globin synthetic ratio has been obtained, but apparent exceptions have been encountered. One of these is due to a common but previously undescribed mutation inside the normal *EcoRI* site in the 5' flanking sequence of the α -globin gene.

MATERIALS AND METHODS

Blood samples were obtained with informed consent from White and Black American donors in the Los Angeles and San Diego area. Thirty-seven unrelated Black and 13 unrelated White subjects were studied. Some Blacks were chosen on the basis of their α -globin-to- β -globin synthetic ratios, and the prevalence of the 19,23 genotype cannot, therefore, be considered characteristic of the Black population. The α -globin-to- β -globin synthetic ratio was measured by a published technique (3), using either [³⁵S]methionine or [³H]leucine as a label. Leukocytes were separated from the blood samples, and DNA was purified, digested with the restriction endonuclease, electrophoresed on agarose, transferred to nitrocellulose paper, and hybridized with α -globin gene probes by using described methods (4, 5). The *EcoRI*, *Bam*HI, and *Hind*III restriction endonucleases were obtained from New England BioLabs. cDNA probes for α -globin genes were grown in pMB9 plasmids in *Escherichia coli* provided through the courtesy of Bernard Forget.

RESULTS

The results of restriction analysis of the 37 Black subjects were as follows: 20 had only the normal 23-kb band; 1 had only a 19-kb band; 1 had a 25-kb and a 23-kb band; and 15 had a 23-kb band and a second more rapidly moving band. Careful inspection of the gels revealed, however, that in the case of 3 of the latter individuals the fast-moving (smaller) band was not the 19-kb fragment characteristic of the single α locus deletion, but rather moved somewhat more rapidly, representing a size of approximately 17 kb. The *EcoRI* restriction fragment in such individuals was similar to that produced when a single α gene deletion is present. However, when the distance over which the DNA fragments were electrophoresed was increased by 50–60%, the 23,19 pattern could readily be distinguished from the 23,17 pattern on side-by-side electrophoresis (Fig. 2). All three of these subjects had been found to have normal α -globin-to- β -globin synthetic ratios. Because the synthetic ratios of these subjects were normal we suspected that they did, in reality, have all four normal globin genes and that the aberrant pattern was due to a mutation that resulted in the appearance of a new *EcoRI* restriction site in the "leftward" or 5' flanking sequence. If this were the case, the globin gene-bearing segment should be normal when studied by using the *Bam*HI restriction endonuclease, an enzyme that also cleaves DNA outside of both α loci (see Fig. 1). Mapping of the DNA with *Bam*HI revealed that in each case only the normal α -globin-bearing segment was, indeed, present (Fig. 2). Further evidence that the 23,17 pat-

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.

Abbreviation: kb, kilobase(s).

FIG. 1. Map of the α -globin region of the human genome showing the location of some of the restriction endonuclease sites, modified from Embury *et al.* (1). A region that has suffered the leftward deletion, most common in the Black population, is shown at the top. The site of the mutation described in this paper is indicated with an asterisk.

tern was due to a mutant *EcoRI* restriction site in the leftward flanking sequence was obtained by double digesting with *EcoRI* and *HindIII*. The digest prepared with these restriction endonucleases contained, in addition to the three expected 17-, 3.7-, and 1.8-kb fragments, an additional fragment 12 kb in length (Fig. 3). In one case family studies were performed. The father and one sib were normal, but two other sibs showed the 23,17 pattern.

A normal 23,23 pattern was found in all 13 White subjects studied.

DISCUSSION

In addition to the normal 23,23-kb restriction pattern the 23,19-kb pattern typical of the single α -globin gene deletion is known to be common in the Black population. The much rarer triplicated α -locus is also known to occur (6, 7), and we found one individual with a 25,23-kb restriction pattern presumably presenting this condition. In addition, however, we have identified a previously undescribed, common mutation in the Black population that results in the appearance of a new *EcoRI* restriction site. On treatment with this restriction endonuclease a segment that is 17 kb in length, approximately 6 kb shorter than the

normal 23-kb α -globin gene-bearing *EcoRI* restriction fragment, is formed. Thus, the new site must be approximately 6 kb "inside" the normal site. It could not be located at the 3' end

FIG. 2. *EcoRI* and *BamHI* restriction fragment patterns of human α -globin genes. Lanes 1, heterozygote for the mutant *EcoRI* restriction site; lanes 2, normal control; lanes 3, heterozygote for the single α gene locus deletion; lanes 4, sib of subject whose pattern is shown in lane 1.

FIG. 3. Restriction fragment patterns of human α -globin genes cleaved by both *EcoRI* and *HindIII* endonucleases. Lane 1, subject with the 23,17 *EcoRI* pattern; lane 2, subject with the single α locus deletion (23,19 *EcoRI* pattern); lane 3, normal subject (23,23 *EcoRI* pattern). Patterns 2 and 3 have an identical appearance because the normal chromosome in the individual with a single α locus deletion provides all of the normal DNA bands. In the case of the individual with the 23,17 *EcoRI* pattern (lane 1) a band representing the portion between the mutant *EcoRI* site and the left intragenic *HindIII* site is seen. At approximately 12 kb in length, this fragment is shorter than the normal 17-kb fragment by the distance between the leftward *HindIII* site and the mutant *EcoRI* site.

of the structural gene complex, because the normal *EcoRI* site at this location is only 1.5 kb from the end of the structural gene. The mutation must therefore be located close to the normal *Xba* I site located approximately 11 kb from the 5' end of the structural gene complex (see Fig. 1). This location was confirmed by double digestion with *EcoRI* and *HindIII* restriction endonucleases.

The function of the long DNA sequences flanking structural genes is not yet clear. Although these regions are relatively highly conserved, so that the restriction map has a relatively uniform pattern from individual to individual, a number of polymorphisms involving these sequences have been described (8–10). Closely linked to structural gene mutations, such polymorphisms may be applicable in prenatal genetic analysis and may also be useful in population studies. The mutation that we report here seems to be highly prevalent in the Black population. The technique of restriction mapping is a complex one, and a moderate degree of variation in the apparent length of gene-bearing segments is inevitable because of technical factors. Although, as shown in Fig. 2, the 17-kb fragment obtained by the mutation that we have described in this report is clearly differentiated from the 19-kb fragment obtained in the common Black form of α -thalassemia, the 17,23 pattern could readily be confused with the 19,23 pattern in shorter electrophoretic runs, especially in the absence of excellent results with side-to-side comparison. It is desirable, therefore, to confirm the diagnosis

of a single gene deletion by using the *EcoRI* map with a second restriction endonuclease such as *BamHI*.

The technology of DNA mapping with restriction endonucleases is sufficiently complex so that it has been applied to only a relatively limited human population. The flanking sequences may prove to be much more permissive than are the structural sequences and, with the number of restriction endonucleases now available, an enormous degree of variability is to be expected in human populations.

This work was supported by Grant HL25552 from the National Heart, Lung, and Blood Institute.

1. Embury, S. H., Miller, J. A., Dozy, A. M., Kan, Y. W., Chan, V. & Todd, D. (1980) *J. Clin. Invest.* **66**, 1319–1325.
2. Dozy, A. M., Kan, Y. W., Embury, S. H., Mentzer, W. C. & Wang, W. C. (1979) *Nature (London)* **280**, 605–607.
3. Tegos, C. & Beutler, E. (1980) *Clin. Lab. Haematol.* **2**, 191–197.
4. Embury, S. H., Lebo, R. V., Dozy, A. M. & Kan, Y. W. (1979) *J. Clin. Invest.* **63**, 1307–1310.
5. Southern, E. (1979) *Methods Enzymol.* **68**, 152–176.
6. Higgs, D. R., Old, J. M., Pressley, L., Clegg, J. B. & Weatherall, D. J. (1980) *Nature (London)* **284**, 632–635.
7. Goossens, M., Dozy, A. M., Embury, S. H., Zachariades, Z., Hadjiminis, M. G., Stamatoyannopoulos, G. & Kan, Y. W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 518–521.
8. Kan, Y. W. & Dozy, A. M. (1978) *Lancet* **ii**, 910–912.
9. Kan, Y. W. & Dozy, A. M. (1980) *Science* **209**, 388–391.
10. Kan, Y. W., Lee, K. Y., Furbetta, M., Angius, A. & Cad, A. (1980) *N. Engl. J. Med.* **302**, 185–188.