Sequence correction and identification of a Rsa I site 3' to human G γ globin gene

Angela Ragusa, Dominique Labie and Rajagopal Krishnamoorthy

Institut de Pathologie Moléculaire, INSERM U.15, CHU Cochin-Port-Royal, 24 rue du Fg St Jacques, 75014 Paris, France
Submitted April 11, 1987

The nucleotide sequence of human DNA encompassing the duplicated γ globin genes has been published (1) and computer derived restriction map is available in the databank. Due to the high degree of sequence homology between the duplicated units only few differences in restriction sites are observed between these two blocks. One such variation is recorded with Rsa I which should produce discriminative 3.44 kb G γ and 2.15 kb A γ specific fragments. In a study of γ genes from thalassemic individuals cloned in our laboratory, we encountered a Rsa I restriction map incompatible with the published data. We always obtained a single γ specific fragment of 2.15 kb albeit the presence of both G γ and A γ genes in our clones (as verified by other digestions). Since a similar pattern was obtained with uncloned genomic counterparts of these cloned fragments, the possibility of changes during cloning in *E. Coli* was eliminated. An alternative is the existence of polymorphism for this restriction enzyme. Study of genomic DNA from individuals of various populations (Mediterraneans, black Africans and Indians) eliminated such a possibility. Patterns of double digestions with Bam HI and Rsa I on the one hand and Pst I and Rsa I on the other hand allowed us to locate the missing Rsa I site in the published sequence to 3' of G γ gene at a position homologous to that of A γ gene. Sequence inspection of the cloned fragment determined this position to 196 nucleotide downstream to the poly(A) addition signal of G γ gene (equivalent position in A γ gene is 189 nt downstream to the signal). Hence the original sequence 8820 ATAC 8825 must be corrected to ---GTAC---. The corrected restriction site is of particular interest to isolate the identical blocks of G γ and A γ fragments encompassing the whole individual genes with their regulatory elements (400 nt 5' to the cap site to $\approx$ 200 nt 3' to the poly(A) addition signal) for subcloning and expression studies. In addition, this enzyme recognizing a 4 bp sequence is employed at present in our laboratory for enriching the genomic γ gene fragments for direct cloning in plasmid libraries (2).

REFERENCES

1. Shen, S., Slighon, J.L., and Smithies, O. (1981) *Cell* 26, 191-203.
2. Nicholls, R.D., Hill, A.V.S., Clegg, J.B., and Higgs, D.R. (1985) *Nucl. Acid Res.* 13, 7569-7578.