Volume 5 Number 2 February 1978 Nucleic Acids Research

Insertion of synthetic copies of human globin genes into bacterial plasmids

J.T. Wilson, L.B. Wilson, J.K. deRiel, L. Villa-Komaroff*, A. Efstratiadis*, B.G. Forget and S.M. Weissman

Departments of Human Genetics and Internal Medicine, Yale University School of Medicine, New Haven, CT 06510 and *Biological Laboratories, Harvard University, Cambridge, MA 02138, USA

Received 22 December 1977

ABSTRACT

Double stranded human globin cDNA was synthesized by use of viral reverse transcriptase from globin mRNA of cord blood of premature infants requiring exchange transfusions. The cDNA was introduced into plasmids and the recombinant DNA plasmids used to transform E. coli X1776. A number of transformants were obtained. Plasmid DNA from selected colonies was isolated and characterized for the type of globin cDNA it contained by three types of procedures: 1) hybridization to previously characterized ^{3}H-labeled α,β and γ cDNA; 2) analysis of the size and nature of fragments produced by digestion of the plasmid DNA by different restriction endonucleases; and 3) by rapid DNA sequence analysis of selected DNA fragments produced by restriction endonuclease digestion. Analysis by these techniques of plasmid DNA from different colonies has definitively identified the presence of human α, β or γ cDNA sequences in different plasmids.

INTRODUCTION

The family of genes encoding the globin polypeptide chains of human hemoglobin includes one β chain gene, one δ chain gene, two α chain genes and two γ chain genes, in addition to at least one gene each for the two embryonic (ε and ζ) globin chains (reviewed in Ref. 1). The globin genes are one of the best characterized eucaryotic gene systems available to study problems of gene regulation and the process of differentiation. Extensive structural information is already available for the human β globin chain mRNA and partial information is available for the human α chain mRNA (2-6). There are limitations in performing extensive structural analyses of isolated mRNAs both because of possible heterogeneity (polymorphism) between multiple gene loci for the same globin chain (7), and because of the difficulty in obtaining large quantities of pure mRNA. To facilitate the nucleotide sequence determination of the α and γ globin chain structural genes and to provide pure molecular probes for the human globin genes which can be used in mapping experiments of normal and mutant (thalassemic) globin genomic DNA, we have prepared cloned bacterial plasmids

containing synthetic DNA copies (cDNA) of human α, β or γ globin gene sequences. The present report describes the characterization of both recombinant plasmids previously reported (7) and newly obtained plasmids.

MATERIALS AND METHODS

(a) Enzymes

Restriction endonucleases Eco RI, Bam HI, HhaI, AluI and Mbo II were obtained from New England BioLabs. RNA-dependent DNA polymerase of avian myeloblastosis virus was provided by Drs. D. and J.W. Beard through the Office of Program Resources and Logistics, Viral Cancer Program, Viral Oncology, Division of Cancer Cause and Prevention, National Cancer Institute. Calf thymus deoxynucleotidyl terminal transferase was a gift of Dr. R. Ratliff. Other enzymes were the same as previously described (3).

(b) Bacterial Strains, Plasmids and Media

Escherichia coli (E. coli) strain HB101 carrying the pMB9 plasmid was obtained from Dr. H. Boyer. The pMB9 plasmid has a molecular weight of 3.5×10^6 and confers tetracycline resistance to the bacterial host cell. This plasmid has a single Eco RI and a single Bam HI restriction endonuclease site (8). E. coli strain HB101 carrying the pCR1 plasmid was obtained from Dr. D. Helinski.

Transformation experiments with hybrid DNAs utilized the EK2 certified recipient cell strain E. coli X1776, obtained from Dr. Roy Curtiss III (9). The rabbit β globin cDNA recombinant DNA plasmid PβG-1 used in certain experiments is the same as previously described (10).

E. coli HB101 containing plasmid DNA was grown in enriched medium (Luria broth) containing, per liter: 10g of trypone; 5g of yeast extract (Difco); 5g of NaCl; 1g of glucose; 1 ml of 1N NaOH, and 25 mg of tetracycline (for pMB9) or 10 mg of kanamycin (for pCR1). E. coli X1776 was grown in the same medium (with or without the above antibiotics) supplemented with 100 mg/ml of diaminopimelic acid, 40 mg/ml of thymidine, 1mg/ml of biotin and 25 mg/ml of nalidixic acid according to Curtiss et al. (9). E. coli X1776 was also grown on agar plates containing the same medium.

(c) Preparation of Plasmid DNA

Plasmid DNA was isolated by the method of Sidikaro et al. (11) as follows: E. coli containing plasmid was grown at 37° in Luria broth to middle log phase. Chloramphenicol (150 μg/ml) was added and the cells were allowed to continue incubation with vigorous agitation at 37°C for 18-24 hours. Cells were harvested and lysates were prepared by the method of

Guerry et al (12). Crude lysates were adjusted with cesium chloride to a density of 1.395 and centrifuged in a Beckman type 40 rotor at 40,000 rpm for 48 hours at 15°C. The plasmid DNA bands, stained with ethidium bromide, were collected, extracted with iso-amyl alcohol, dialysed against 0.3M sodium acetate and precipitated by the addition of 2 volumes of ethanol. The DNA was recovered by centrifugation at 9000 rpm for 10 minutes.

(d) Globin mRNA Isolation

Total reticulocyte RNA was isolated from cells of adults or premature infants (requiring exchange transfusion for neonatal jaundice usually due to Rh blood group incompatibility) by phenol extraction of total red cell lysates as described (13). Globin mRNA was purified from the total RNA by a two step procedure: sucrose gradient centrifugation followed by oligo (dT)-cellulose column chromatography of the RNA sedimenting between 5s and 16s (14,15). Human globin mRNA purified in this manner gives a single broad band (of approximately 200,000 M.W.) without any detectable higher molecular weight RNA when analysed by polyacrylamide gel electrophoresis in non denaturing conditions (15). When analysed by formamide-polyacrylamide gel electrophoresis, the RNA is resolved into two discrete bands corresponding to the α and β mRNAs, in the case of human adult globin mRNA, and a third band (γ mRNA) is found between α and β mRNA in the case of globin mRNA from fetal or cord blood reticulocytes (14,16). When mRNA prepared in this manner is labeled in vitro with ^{125}I and subjected to fingerprint analysis, it gives a pattern consistent with a homogeneous RNA species containing essentially only α and β globin mRNA sequences (15); in the case of cord blood mRNA, γ mRNA sequences are also seen. Globin mRNA purified in this manner is at least 95% pure. Globin cDNA synthesized from mRNA of premature newborn cord blood cells was approved for cloning under P3 + EK2 conditions of containment by our institutional biohazard committee and by the NIH Recombinant DNA Molecule Program Advisory Committee, under the provisions of the NIH Recombinant DNA Research Guidelines which allow lowering of containment by one level in the case of embryonic primate material because of the very low probability that such material was ever exposed to horizontally transmitted viruses (17).

(e) cDNA Synthesis

cDNA was synthesized from the purified cord blood globin mRNA using reverse transcriptase as previously described (18-20). With single stranded globin cDNA as template, second strand synthesis was accomplished in a

self primed reaction (20,21) using either E. coli DNA polymerase I (20) or reverse transcriptase (7). The resulting double stranded "hairpin" cDNA molecules were digested with S_1 nuclease as described (7,20) to cleave the hairpin structure and convert the cDNA to an open double stranded linear molecule. In some cases, the double stranded cDNA was fractionated by polyacrylamide gel electrophoresis (10) to select the largest ("full-length") double stranded cDNA molecules for construction of recombinant DNA plasmids.

(f) Construction of Recombinant DNA Plasmids

Recombinant DNA molecules consisting of plasmid + human globin cDNA were constructed by the dA-dT tailing procedure that has been described previously (10,21). For our studies, we utilized the procedure of Roychoudury et al. (22) for the addition of homopolymer dT tails to plasmid DNA (after cleavage with Eco RI) and dA tails to the globin cDNA. After the addition of the homopolymer tails, equimolar amounts of the two molecules (at a concentration of 0.35 μg/ml of plasmid DNA) were annealed for 3 min at 60°C in 0.1 M NaCl, 10mM Tris-HCl, pH 7.5, and 0.2 mM EDTA. After the 60°C incubation, the molecules were further annealed for 3 hr. at 42°C and then slowly allowed to cool to room temperature. Other recombinant DNA plasmids were constructed by "blunt-end" ligation as previously described (7). 20 ng of the annealed recombinant DNA were used for each individual transformation experiment.

(g) Transformation of X1776 by Recombinant DNA Plasmids

E. coli X1776 was transformed by recombinant DNA plasmids according to the procedure of Curtiss et al. (9). Two ml of an overnight culture of X1776 were diluted into 20 ml of fresh medium and the bacteria were grown at 37°C with shaking until the cell concentration reached 1.5×10^8 bacteria per ml. The cells were then sedimented by centrifugation at 8000 rpm and the cell pellet gently resuspended in 10 ml of 10 mM NaCl. The cells were again sedimented as above and gently resuspended in buffer containing 75mM $CaCl_2$, 10mM Tris-HCl, pH8.0 and 0.8% NaCl. The bacteria were left at room temperature for 20 minutes and again sedimented as described above. The cell pellet was then resuspended in 2.0ml of the same buffer and 0.2ml of the suspension added to 0.1 ml of plasmid DNA (0.2 μg/ml).The mixture was kept on ice for 25 minutes, rapidly heated to 42°C in a water bath for one minute then replaced at 4°C for 10 minutes. One ml of X1776 culture medium was then added to the suspension. After incubation at 37° for 1 hr, the cell suspension was applied to agar plates containing X1776 growth medium and incubated at 37°C to test for survival of the bacterial cells

which was usually found to be approximately 4%. Duplicate plates containing the appropriate antibiotics were cultured to select for transformants.

(h) Colony Hybridization and Isolation of Cloned Recombinant DNA Plasmids

Colonies of transformed bacteria which grew in the presence of antibiotics were screened for the presence of globin cDNA sequences by colony hybridization according to the procedure of Grunstein and Hogness (23). Bacteria from individual colonies were transferred from agar plates to 8cm nitrocellulose filters by contact, then lysed and their DNA denatured in situ (23). Human globin mRNA was used as hybridization probe after being labeled in vitro with ^{32}P by the following method. The mRNA was partially hydrolyzed in 0.1M Tris-HCl, pH 9.5, for 2 min. at 85°C (24), then the 5'-termini of the RNA fragments were labeled using [γ-^{32}P] ATP (1000 to 4000 Ci/m mole, ICN or New England Nuclear) and polynucleotide kinase as described (24). After labeling, the mRNA was purified by passage over a column of Sephadex G25 and precipitated with alcohol. $1x10^5$ cpm of the labeled mRNA were hybridized to each nitrocellulose filter in 5x SSC and 50% formamide for 24 hours at 37°. After hybridization, the filters were extensively washed with 1X SSC, treated with pancreatic RNase (20µg/ml) dried and subjected to autoradiography. Colonies which gave positive hybridization were picked from replica plates, grown individually in suspension cultures and plasmid DNA isolated as described above.

(i) Gel Electrophoresis and Restriction Endonuclease Digestion

DNA was digested with restriction endonucleases AluI, Bam HI, or HhaI at 37°C in 6 mM Tris HCl, pH 7.5, 6 mM $MgCl_2$, 50 mM NaCl, and 6 mM β mercaptoethanol. Digestions with Eco RI were carried out at 37°C in 100 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, and 50 mM NaCl. Digestions with MboII were performed at 37°C in 6 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 6 mM NaCl, and 6 mM β mercaptoethanol. The number of enzyme units used and the time of digestion varied with different batches of enzyme.

Native or digested plasmid DNAs were fractionated by electrophoresis in slab gels of 1% agarose or 7% polyacrylamide. The gels were stained with 10µg/ml ethidium bromide to visualize the DNA bands. For hybridization experiments, DNA was transfered from agarose gels to nitrocellulose filters by the gel blotting technique described by Southern (29) and hybridized with ^{32}P-labeled mRNA.

(j) Hybridization Assays

Single stranded human α and β specific cDNAs were synthesized in the presence of ^3H dCTP (25 Ci/m mole) as described (14,19) and used as hybrid-

ization probes. ^{3}H-labeled human γ cDNA was purified from mixed α,β,γ cDNA (synthesized from total cord blood globin mRNA) by hybridization of the cDNA to adult ($\alpha+\beta$) globin mRNA followed by hydroxylappatite column chromatography and isolation of the non-hybridized (γ) cDNA (25). Non radioactive plasmid DNA was hybridized to the different ^{3}H-labeled cDNAs in 0.2M phosphate buffer, pH 6.8, and 0.5% SDS for 16 hours at 78° and the percentage of the ^{3}H-labeled cDNA hybridized to plasmid DNA was determined by S_1 nuclease digestion (14,19,26).

(k) DNA Sequence Analysis

DNA sequencing of inserted fragments was performed by use of the method of Maxam and Gilbert (27). For these studies, recombinant DNA plasmids were digested with restriction endonuclease HhaI. The resulting products were separated by electrophoresis in a vertical 40 cm long slab gel apparatus as described by Subramanian et al. (28) except that the upper 2/3 of the gel consisted of 6% acrylamide and the lower 1/3, 10% acrylamide. After electrophoresis, the gel was stained with ethidium bromide and the DNA fragment that contained the inserted cDNA sequence (as determined by its different mobility compared to that of the control plasmid DNA fragments) was excised. The gel segment was ground in 0.1X SSC pH 7.5 in a Tekman tissue homogenizer. The suspension was then repeatedly centrifuged in a conical glass tube to remove acrylamide. DNA was precipitated by the addition of 2 volumes of cold ethanol. After storage at -70° for 1 hr., the DNA was recovered by centrifugation and dried under vacuum. The isolated fragment was then labeled with ^{32}P at its 5'-termini using bacterial alkaline phosphatase, [γ^{32}P] ATP and polynucleotide kinase as described (3). After labeling, the fragment was recut by a second restriction endonuclease and the resulting products separated again by gel electrophoresis, excised, eluted and sequenced.

RESULTS AND DISCUSSION

(a) Transformation Experiments

Transformants of E. coli X1776 by recombinant pMB9 cDNA plasmids yielded 5 to 9×10^{4} transformed colonies per μg of plasmid DNA. Similar experiments with pCR1-cDNA recombinant DNA plasmids produced a 10 fold lower yield of transformed colonies. This reduction in the frequency of transformation using pCR1 probably represents a loss in efficiency due to the increased molecular weight of the pCR1 plasmid, as compared to pMB9 (30).

(b) Colony Hybridization

E. coli X1776 clones that acquired antibiotic resistance were picked

and tested for hybridization of their DNA to ^{32}P-labeled human globin mRNA by the technique of colony hybridization (23). Figure 1 shows the resulting autoradiograph from one such colony hybridization study: several colonies show strong positive hybridization with the human globin mRNA probe (colonies A to E), whereas many colonies show no significant hybridization. In a typical experiment only about 2-8% of the colonies gave positive hybridization. This result is most probably due to the preferential transformation of X1776 by uncut plasmid DNA, over that by recombinant plasmid DNA molecules. As a positive control, the rabbit β globin cDNA recombinant plasmid PβG1 of Maniatis et al (10) [Colony R in Fig. 1] was included in our study and gave weaker but significant hybridization with the ^{32}P-labeled human globin mRNA as would be expected from the sequence homology between human and rabbit β globin mRNAs (31). More than 50 positive colonies have thus far been isolated.

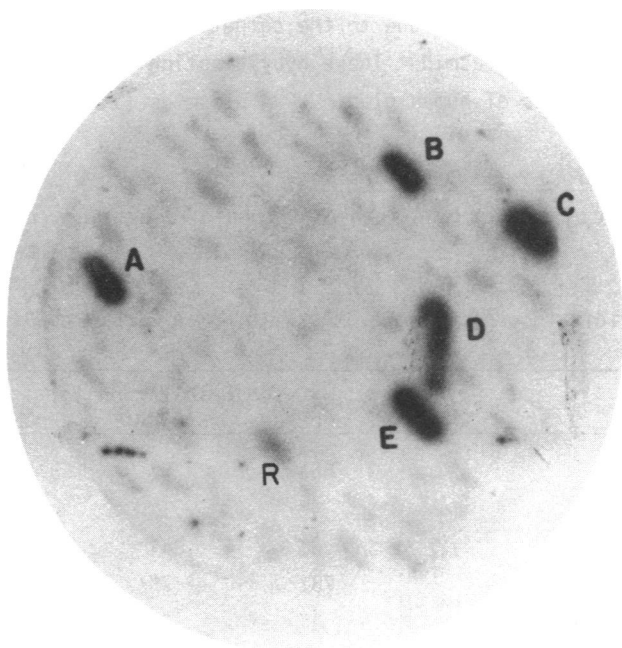

Figure 1. Autoradiograph of hybridization of ^{32}P-labeled human globin mRNA with DNA of colonies of E. coli X1776 transformed by recombinant DNA plasmids consisting of pMB9 and human fetal globin cDNA. In situ hybridization of DNA fixed to nitrocellulose filters was performed using ^{32}P-labeled human adult globin mRNA prepared as described in Methods. A to E indicate transformed colonies which hybridize to the globin mRNA. R indicates a colony transformed by the rabbit β globin cDNA plasmid PβG1.

(c) DNA-cDNA Hybridization Results

DNA isolated from a number of our plasmids was hybridized in liquid hybridization assays, with purified ^{3}H-labeled human α, β and γ globin cDNA and the results are listed in Table 1. DNA of recombinant plasmids JW101 and JW112 gave 65% and 73% hybridization respectively with labeled α globin cDNA but only 6% and 7% hybridization with β globin cDNA. DNA of plasmids JW102 and JW103 showed opposite results. They gave only 4% to 5% hybridization with α globin cDNA but 51% and 84% hybridization respectively with β cDNA. Neither DNA hybridized with γ specific cDNA. DNA of plasmids JW151 and JW152 gave 40% and 55% hybridization respectively with human γ cDNA but insignificant hybridization with α and β cDNAs. Control pMB9 DNA gave no significant hybridization with either α or β globin cDNA probes. The low level of hybridization of β cDNA-plasmid DNA with α cDNA (and vice versa) is probably due to the 10 to 15% contamination of the β cDNA probe with α cDNA sequences (and vice versa). The failure of the plasmid DNA to hybridize totally to the cDNA is probably due to the competitive hybridization by the second DNA strand of the plasmid. These hybridization results indicate that all three expected types of human globin cDNA plasmids (α, β and γ) have been isolated.

(d) Gel Electrophoresis and Gel Blotting

The insertion of globin cDNA into pMB9 should result in a recombinant

Table 1

Hybridization of Plasmid DNAs With Human Globin cDNAs

Plasmid DNA	% Hybridization to: α cDNA	β cDNA	γ cDNA
JW101	65%	6%	0
JW102	4%	51%	0
JW103	4.5%	84%	0
JW112	73%	7%	0
JW151	4%	3%	40%
JW152	7%	3%	55%
pMB9	0	0	

Plasmid DNA was hybridized with 500 cpm of ^{3}H-labeled chain specific human globin cDNA in 0.2M sodium phosphate, pH 6.8, and 0.5% SDS for 16 hours at 78°; the percentage of hybridization was determined by S1 nuclease digestion after dilution of the 10 μl reaction into 2 ml of S1 digestion buffer (26).

DNA plasmid that shows a difference in mobility compared to native pMB9 DNA after agarose gel electrophoresis. Various purified recombinant plasmid DNAs were therefore compared to native pMB9 DNA with respect to electrophoretic mobility in agarose as well as ability to hybridize with ^{32}P-labeled human globin mRNA by use of the gel blotting technique described by Southern (29), and the results are shown in Figure 2.

The rabbit β cDNA plasmid (lane A) and three of our human globin cDNA plasmids, JW101 (lane C), JW102 (lane D) and JW103 (lane F) all show a similar electrophoretic mobility (Fig. 2a) which is somewhat slower than that of native pMB9 plasmid DNA (lane B) and plasmid DNA (lane E) from one of the X1776 transformants that gave negative colony hybridization with labeled human globin mRNA. Similar results were obtained with plasmid DNA obtained from a number of colonies giving positive colony hybridization (data not shown). The electrophoretic mobility of the human globin cDNA recombinant plasmids, being similar to that of PβG1, therefore suggests that they contain approximately 800 additional base pairs of DNA (including A-T tails) compared to native pMB9 plasmid DNA.

DNA from the agarose gel shown in Fig. 2a was then transferred to a nitrocellulose filter by blotting and hybridized with ^{32}P-labeled human globin mRNA, as shown in Fig. 2b. The rabbit β cDNA plasmid (lane A) gives a low level of hybridization whereas human globin cDNA plasmids JW101 (lane C), JW102 (lane D) and JW103 (lane F) give strongly positive hybridization with the labeled human globin mRNA probe. Control pMB9 DNA (lane B) and plasmid DNA from the negative X1776 transformant (lane E) show no significant hybridization with the probe. These results confirm those of the colony hybridization study and indicate that recombinant DNA plasmids JW101, JW102 & JW103 contain inserts of human globin cDNA.

(e) Restriction Endonuclease Digestion of Recombinant DNA Plasmids

Alteration of the normal restriction endonuclease digestion pattern of pMB9 DNA is expected with the insertion of a cDNA fragment into the plasmid and the new pattern should help identify the sequence specificity of the cDNA insert. Three clones (JW101, JW102, JW103) were chosen for study from the clones with plasmid DNA that gave positive hybridization and showed larger molecular size than pMB9 DNA after agarose gel electrophoresis. Purified DNAs from these plasmids and pMB9 were digested with restriction endonucleases Eco RI or Bam HI and the undigested plasmid DNAs (Figure 3, lanes A to D) as well as the DNA digests (Figure 3, lanes E to L) were fractionated by electrophoresis in a 1% agarose gel. The restriction endonuclease map

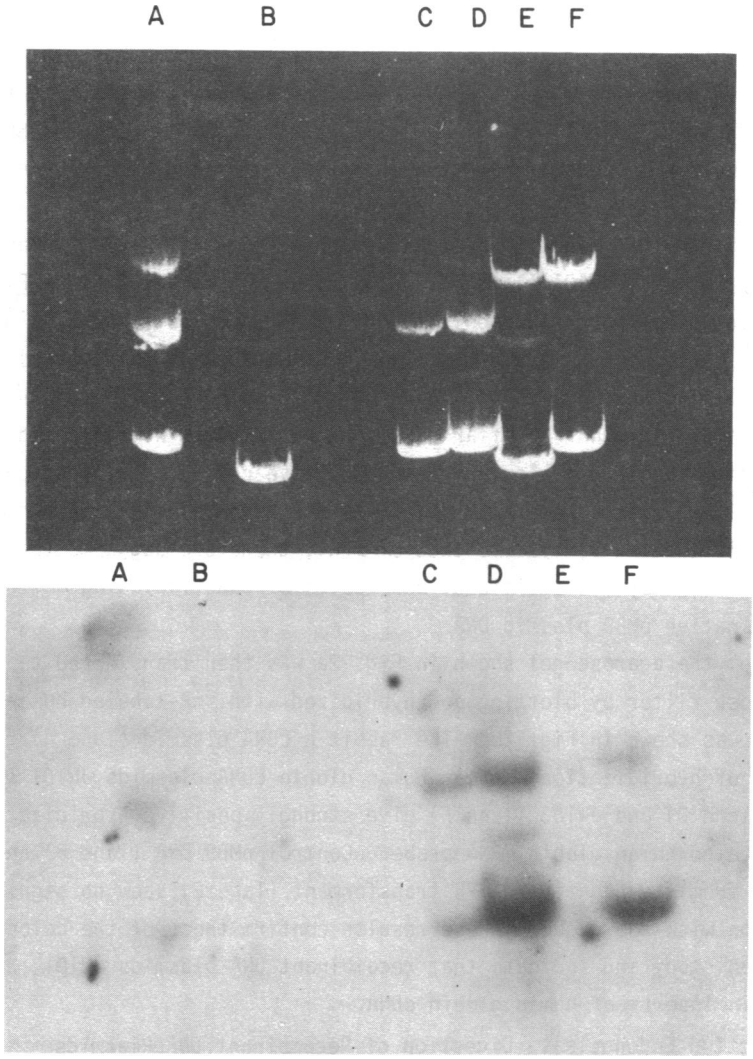

Figure 2. Agarose gel electrophoresis of plasmid DNAs. a) Visualization of DNA by staining with ethidium bromide. One to three DNA bands are visualized in different preparations: the lower more rapidly migrating band represents form I plasmid DNA; the band of intermediate mobility represents form II plasmid DNA; the upper less rapidly migrating band (in lanes A, E and F) represents E. coli host DNA. Lane A: rabbit β globin cDNA plasmid PβG1; lane B: pMB9 plasmid DNA; lane C: human globin cDNA plasmid JW101; lane D; human globin cDNA plasmid JW102; lane E: Plasmid DNA from an X1776 transformant giving negative colony hybridization; lane F: human globin cDNA plasmid JW103. b) Autoradiograph of hybridization of plasmid DNA to ^{32}P-labeled human globin mRNA using the gel blotting technique. DNA from the agarose gel electrophoresis shown in (a) was transfered to a nitrocellulose filter and hybridized to ^{32}P-labeled human adult globin mRNA.

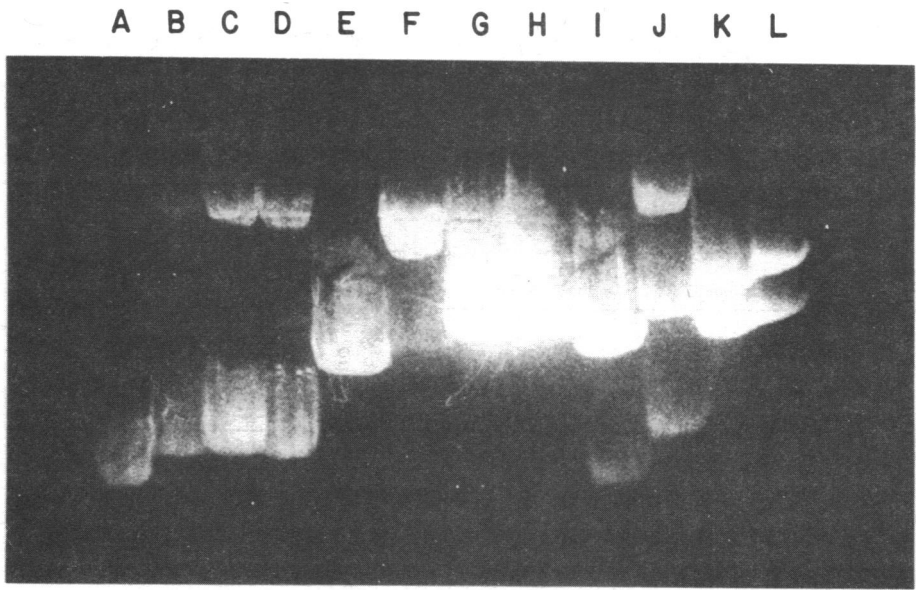

Figure 3. Digestion of plasmid DNAs with restriction endonucleases. Lanes A to D: undigested DNA of pMB9 (A), JW101 (B), JW102 (C), and JW103 (D). Lanes E to H: Eco RI digests of pMB9 (E), JW101 (F), JW102 (G) and JW103 (H). Lanes I to L: Bam HI digests of pMB9 (I), JW101 (J), JW102 (K) and JW103 (L). Fractionation was by electrophoresis in 1% agarose gel.

which we have independently determined for human α and β globin cDNAs (Figure 4),showed that the human β globin structural gene contains a single Eco Rl site (at codons 121-122) and a single Bam HI site (at codons 98-100) (2), whereas the α globin gene contains neither of these sites (Ref. 3 and unpublished). The results shown in Figure 3 demonstrate that Eco RI converted Form I DNA to linear DNA (indicating one cleavage site) in the cases of pMB9 (lane E), JW102 (lane G) and JW103 (lane H), but did not convert Form I DNA to linear DNA in the case of plasmid JW101 (lane F) which must therefore not contain an Eco RI site. Since the recombinant plasmids were produced by the addition of dA - dT tails at the single Eco RI site of pMB9 (thus destroying this site), the Eco RI cleavage site in plasmids JW102 and JW103 must result from the insertion into pMB9 of a cDNA fragment that contains a single Eco RI site. Since plasmid JW101 no longer contains an Eco RI site, the inserted cDNA fragment that it contains does not have an Eco RI site. Digestion of the plasmid DNAs with restriction endonuclease Bam HI (lane I to L) showed that pMB9 DNA (lane I) and plasmid JW101 (lane J) were almost totally converted to linear DNA indicating the presence of a single

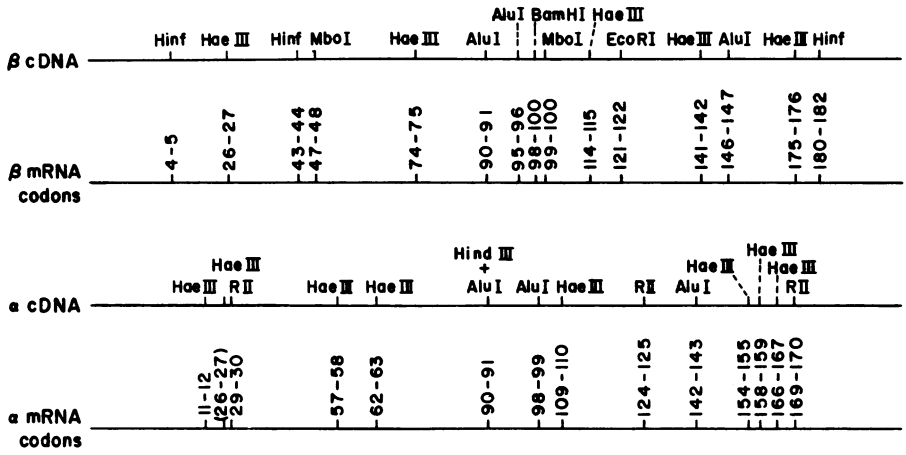

Figure 4. Restriction endonuclease cleavage map of human α and β globin cDNAs (2,3).

Bam site in the DNAs. However, plasmids JW102 (lane K) and JW103 (lane L) yielded two distinct DNA fragments indicating the presence of two Bam sites in these plasmids (see also Figure 5, lane D). Since the Bam site in pMB9 was not involved in the tailing procedure, the single molecule yielded by pMB9 and JW101 after digestion with Bam HI must be attributed to cleavage of these DNAs at the single known Bam site of pMB9. One of the two Bam sites found in JW102 and JW103 is also probably the known pMB9 site, but the other site must have been contributed by the cDNA fragment inserted into pMB9. Therefore, the inserted cDNA fragment in plasmids JW102 and JW103 contained a Bam HI and an Eco RI site (consistent with a β cDNA sequence) whereas the inserted cDNA fragment in plasmid JW101 contained neither of these two sites (consistent with an α cDNA sequence).

Plasmid DNA of pMB9, JW101 and JW102 was also digested with endonuclease Hha I (Figure 5). When the rabbit β globin cDNA plasmid PβG1 was cut with this enzyme, a single new product was obtained when compared to a control pMB9 digest (10). The new product had a slower electrophoretic mobility, and was therefore larger than the corresponding product of native pMB9 DNA; it contained the entire rabbit β globin cDNA insert (10). A similar result was obtained when our putative human β cDNA plasmid JW102 was cut with Hha I (Figure 5, lane A) and compared to a similar digest of control pMB9 DNA (lane C). This result is consistent with the absence of a Hha I site in the human β globin gene (Figure 4) as determined in previous sequence

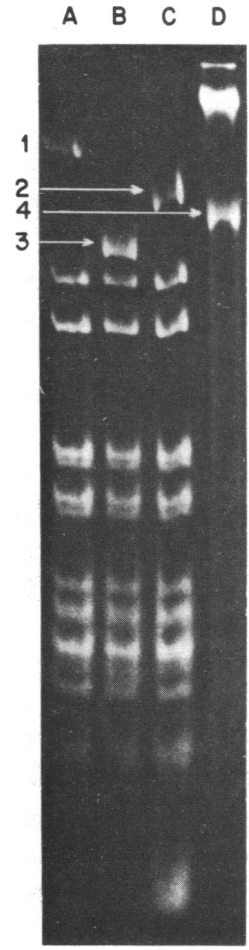

Figure 5. Digestion of plasmid DNAs with restriction endonucleases. Lanes A to C: endonuclease HhaI digests of JW102 (A), JW101 (B) and pMB9 (C). Lane D shows the two restriction fragments that result from Bam HI digestion of JW102. Fractionation was by electrophoresis in 6% polyacrylamide gel.

analysis of human β globin cDNA (2,6). Digestion of plasmid JW101 with Hha I yielded a (? doublet) band (band 3, lane B) which had a more rapid electrophoretic mobility (and therefore a small size) than the altered band 2 of the control pMB9 digest (lane C). The putative α cDNA insert of plasmid JW101 therefore contained a Hha I site, but since only one new band was detected, it is likely that the additional cleavage site produces two fragments of approximately equal size which are not separable by the electrophoretic procedure used. Consistent with this result is our finding from

sequence analysis of de novo synthesized α cDNA, that the α globin structural gene contains a Hha I site within its translated portion (unpublished observations).

Plasmid clones thought to contain γ cDNA inserts were selected for further study in cases where colony hybridization gave negative results with ^{32}P-labeled adult globin mRNA but positive results with ^{32}P-labeled fetal (cord blood) globin mRNA. One such clone was plasmid JW151. Restriction endonuclease digestion of this plasmid with Hha I, Eco RI, Bam HI or HpaII followed by gel electrophoresis, yielded the cleavage pattern characteristic of the putative human β cDNA plasmid JW102 (data not shown). However, cleavage of the putative γ clone JW151 with restriction endonuclease MboII (Figure 6, lane B) revealed the presence of two bands not found in similar digests of native pMB9 DNA (lane D), of the β cDNA plasmid JW102 (lane A) or of the α cDNA plasmid JW112 (lane C). In separate experiments (data not shown) DNA of plasmid JW112 gave restriction endonuclease cleavage patterns similar to those of α cDNA plasmid JW101.

(f) DNA Sequencing Studies

One of our recombinant DNA plasmids (JW102) has been further characterized by determining the nucleotide sequence of a portion of the inserted cDNA fragment. We chose this plasmid, because it contains β cDNA sequences complementary to the human β globin mRNA, and we have recently determined nearly the entire nucleotide sequence of the human β globin mRNA by independent techniques (2). For our sequence analysis, it was advantageous to use a restriction endonuclease that cleaves pMB9 into many fragments but leaves the cDNA insert intact. Digestion of plasmid JW102 with restriction endonuclease Hha I provides such a result (see above).

For sequence analysis, we initially digested JW102 plasmid DNA with restriction endonuclease Hha I and separated the resulting fragments by electrophoresis on a 6% acrylamide gel. The DNA from the top band of such a fractionation (band 1 of Figure 5, lane A) was eluted, recut with Bam HI, and labeled at its 5'-ends with [γ-^{32}P]ATP. The DNA was then recut with endonuclease Mbo II, the fragments separated by gel electrophoresis and subjected to sequence analysis using the technique of Maxam and Gilbert (27). Figure 7 shows a representative autoradiograph of gel separations of DNA products fragmented by use of this technique. The nucleotide sequence which can be "read off" from the autoradiograph is consistent with the sequence of human β globin mRNA (Figure 8) from codon no. 106 to codon no. 127. The inserted cDNA fragment in plasmid JW102 must therefore be a cDNA

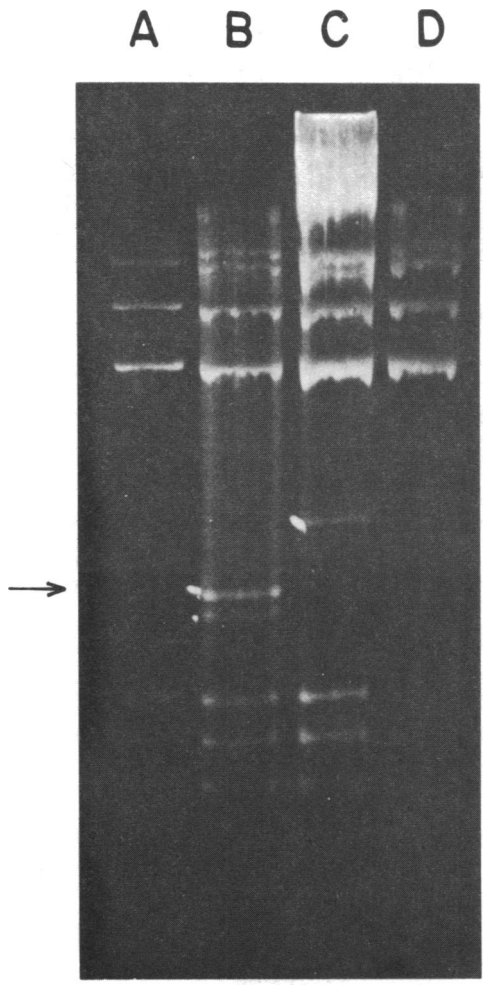

Figure 6. Digestion of plasmid DNAs with restriction endonuclease MboII. Lane A: JW102; Lane B: JW151; Lane C: JW112; Lane D: pMB9. The arrow indicates the position of the additional bands produced by endonuclease digestion of the recombinant γ cDNA plasmid JW151. Fractionation was by electrophoresis in 6% polyacrylamide gel.

copy of human β globin mRNA. From these and other analyses (data not shown), we have determined that this plasmid contains an inserted cDNA fragment that is complementary to the human β globin mRNA approximately from codon no. 20 into the 3'-untranslated portion of the mRNA.

(g) Biohazard Considerations

Cloning experiments and subsequent growth of recombinant DNA plasmids were carried out under conditions of P3 + EK2 containment as specified by the

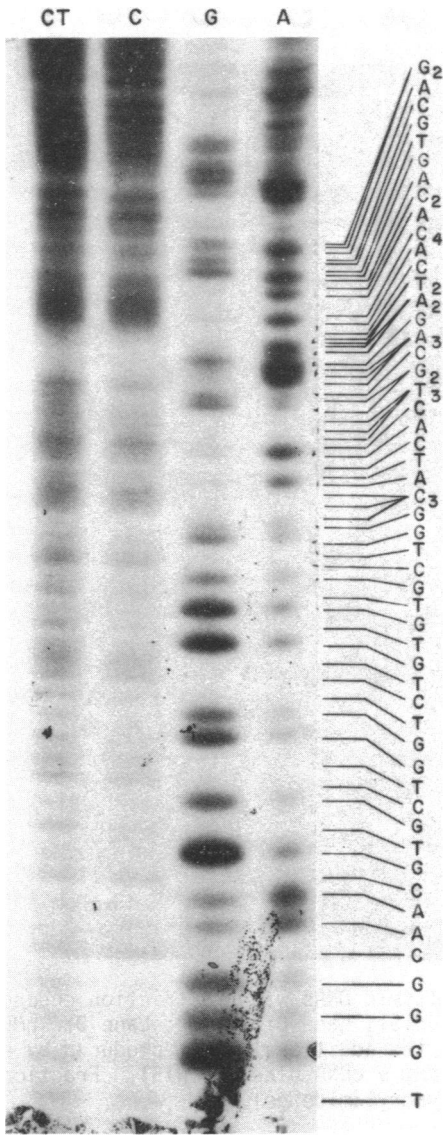

Figure 7. Nucleotide sequence of a DNA fragment of recombinant human globin cDNA plasmid JW102. Sequence analysis was carried out using the method of Maxam and Gilbert (27); the sequence is consistent with that previously determined for human β globin mRNA from codon no. 106 to codon no. 127 (2).

HUMAN β-GLOBIN mRNA SEQUENCE

```
                                0                   5                  10
                                AUG GUG CAC CUG ACU CCU GAG GAG AAG UCU GCC GUU

                  15                  20                  25                  30
ACU GCC CUG UGG GGC AAG GUG AAC GUG GAU GAA GUU GGU GGU GAG GCC CUG GGC AGG

                  35                  40                  45
CUG CUG GUG GUC UAC CCU UGG ACC CAG AGG UUC UUU GAG UCC UUU GGG GAU CUG UCC

 50                  55                  60                  65
ACU CCU GAU GCA GUU AUG GGC AAC CCU AAG GUG AAG GCU CAU GGC AAG AAA GUG CUC

      70                  75                  80                  85
GGU GCC UUU AGU GAU GGC CUG GCU CAC CUG GAC AAC CUC AAG GGC ACC UUU GCC ACA

          90                  95                 100                 105
CUG AGU GAG CUG CAC UGU GAC AAG CUG CAC GUG GAU CCU GAG AAC UUC AGG CUC CUG

                 110                 115                 120                 125
GGC AAC GUG CUG GUC UGU GUG CUG GCC CAU CAC UUU GGC AAA GAA UUC ACC CCA CCA

                 130                 135                 140
GUG CAG GCU GCC UAU CAG AAA GUG GUG GCU GGU GUG GCU AAU GCC CUG GCC CAC AAG

145                 150                 155                 160
UAU CAC UAA GCU CGC UUU CUU GCU GUC CAA UUU CUA UUA AAG GUU CCU UUG UUC CCU

      165                 170                 175                 180
AAG UCC AAC UAC UAA ACU GGG GGA UAU UAU GAA GGG CCU UGA GCA UCU GGA UUC UGC

CUA AUA AAA AAC AUU UAU UUU CAU UGC POLY A
```

Figure 8. Nucleotide sequence of coding and 3'-non coding portions of human β globin mRNA (2,4,6).

NIH Guidelines for Recombinant DNA Research . As discussed in Materials and Methods (Section d), use human fetal (cord blood) globin cDNA was approved for cloning under conditions of P3 + EK2 containment by our institutional biohazard committee and by the NIH Recombinant DNA Molecule Program Advisory Committee. Some cDNA clones were initially obtained (7) under Asilomar guidelines; work with these clones was stopped at the time of issuance of the draft of the NIH Guidelines and DNA from these cDNA plasmids was kept frozen until official adoption of the NIH Guidelines and certification of EK 2 host-vector systems. Cloning experiments were then resumed using X1776 and plasmid pCR1 initially, then pMB9 after its official certification as an EK2 vector. Human β globin cDNA plasmid JW102 has recently been reclassified by the NIH Recombinant DNA Molecule Advisory Committee for work under conditions of P2 + EK1 containment on the basis of the characterization of plasmid JW102 described in this manuscript.

(h) Summary and Conclusions

In summary we have constructed recombinant DNA plasmids using plasmid

DNA and double stranded human globin cDNA synthesized by use of viral reverse transcriptase from human fetal (cord blood) globin mRNA, and individual recombinant globin cDNA plasmids were isolated by bacterial cloning techniques. The cloned recombinant plasmid DNAs were characterized by restriction endonuclease mapping, hybridization to chain specific human globin cDNA and, in one case, by DNA sequence analysis. Our results indicate that the three expected major types of human globin cDNA plasmids (α, β and γ) have been successfully isolated. The availability of these cloned human globin cDNA plasmids will make possible the determination of the nucleotide sequence of unique copies of the duplicated and possibly polymorphic human α and γ globin chain structural loci. The availability of pure chain specific human globin cDNA plasmids should also greatly facilitate the performance of a number of gene mapping studies related to the molecular genetics of hemoglobin synthesis in the normal and in states of abnormal globin gene expression such as the thalassemias.

ACKNOWLEDGEMENTS

This work was supported by grants GM 20124 and CA 05186 of the NIH. J.K. deRiel is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research. L. Villa-Komaroff was supported by a fellowship of the Helen Hay Whitney Foundation and A. Efstratiadis by a fellowship of the Harvard Society of Fellows.

REFERENCES

1. Bunn, H.F., Forget, B.G. and Ranney, H.M. (1977) Human Hemoglobins, W.B. Saunders, Phila. pp. 101-112.
2. Marotta, C.A., Wilson, J.T., Forget, B.G. and Weissman, S.M. (1977) J. Biol. Chem. 252,5040-5053.
3. Wilson, J.T., deRiel, J.K., Forget, B.G., Marotta, C.A., and Weissman, S. M. (1977) Nucl. Acid Res. 4;2353-2368.
4. Proudfoot, N.J. and Longley, J.I. (1976) Cell 9, 733-746.
5. Baralle, F.E. (1977) Cell (in press).
6. Chang, J.C., Temple, G.F., Poon, R.F., Neumann, K.H. and Kan, Y.W. (1977) Proc. Natl. Acad. Sci. USA 74, 5145-5149.
7. Wilson, J.T., Forget, B.G., Wilson, L.B. and Weissman, S.M. (1977) Science 196, 200-202.
8. Rodriguez, R.L., Bolivar, F., Goodman, H.M., Boyer, H.W. and Betlach, M. (1976) in ICN/UCLA Symposium on Molecular Mechanisms in the Control of Gene Expression. D.P. Nierlich, W.J. Rutter and C.F. Fox eds., Academic Press, New York, pp. 471-477.
9. Curtiss, R., III, Pereira, D.A., Hsu, J.C., Hull, S.C., Clark, J.E., Maturin, L.J., Sr., Goldschmidt, R., Moody, R., Inoue, M. and Alexander, L. (1977) in Recombinant Molecules: Impact on Science and Society, R.F. Beers Jr. and E.G. Basset, eds., Raven Press, New York, pp. 45-56.
10. Maniatis, T., Kee, S.G., Efstratiadis, A. and Kafatos, F.C. (1976) Cell 8, 163-182.
11. Sidikaro, J. and Nomura, M. (1975) J. Biol. Chem. 250, 1123-1131.

12. Guerry, P., LeBlanc, D.J. and Falkow, S. (1973) J. Bacteriol. 116, 1064-1066.
13. Benz, E.J., Jr., Swerdlow, P.S. and Forget, B.G. (1975) Blood 45, 1-10.
14. Forget, B.G., Housman, D., Benz, E.J. Jr. and McCaffrey, R.P. (1975) Proc. Natl. Acad. Sci. USA 72, 984-988.
15. Cohen-Solal, M., Forget, B.G., Prensky, W., Marotta, C.A. and Weissman, S. M. (1977) J. Biol. Chem. 252, 5032-5039.
16. Kazazian, H.H. Jr., Ginder, G.D., Snyder, P.G., van Beneden, R.J. and Woodhead, A.P. (1975) Proc. Natl. Acad. Sci. USA 72, 567-571.
17. in NIH Recombinant DNA Research Guidelines (1976) Federal Register 41, 27917.
18. Efstratiadis, A., Maniatis, T., Kafatos, F.C., Jeffrey, A. and Vournakis, J. (1975) Cell 4, 367-378.
19. Forget, B.G., Hillman, D.G., Lazarus, H., Barell, E.F., Benz, E.J., Jr., Caskey, C.T., Huisman, T.H.J., Schroeder, W.A. and Housman, D. (1976) Cell 7, 323-329.
20. Efstratiadis, A., Kafatos, F.C., Maxam, A.M., and Maniatis, T. (1976) Cell 7, 279-288.
21. Higuchi, R., Paddock, G.V., Wall, R. and Salser, W. (1976) Proc. Natl. Acad. Sci. USA 73, 3146-3150.
22. Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucl. Acid. Res. 3, 101-116.
23. Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
24. Maizels, N. (1976) Cell 9, 431-438.
25. Lanyon, W.G., Ottolenghi, S. and Williamson, R. (1975) Proc. Natl. Acad. Sci. USA 72, 258-262.
26. Housman, D., Forget, B.G., Skoultchi, A., and Benz, E.J., Jr. (1973) Proc. Natl. Acad. Sci. USA 70, 1809-1813.
27. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
28. Subramanian, K.N., Pan, J., Zain, S. and Weissman, S.M. (1974) Nucl. Acid. Res. 1, 727-752.
29. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
30. Covey, C., Richardson, D. and Carbon, J. (1976) Mol. Gen. Genet. 145, 155-158.
31. Kafatos, F.C., Efstratiadis, A., Forget, B.G. and Weissman (1977) Proc. Natl. Acad. Sci. USA (in press).