Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 780–784, January 1996
Biochemistry

Z-DNA-forming sites within the human β-globin gene cluster

(negative supercoiling/transcription rate/erythropoietic cells/topological domain)

VERENA MÜLLER*, MIEKO TAKEYA*, SUSANNE BRENDEL*, BURGHARDT WITTIG*, AND ALEXANDER RICH†

*Abteilung Molekularbiologie und BioInformatik, Institut für Molekularbiologie und Biochemie, Freie Universität Berlin, D-14195 Berlin, Germany; and †Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139-4307

Contributed by Alexander Rich, October 2, 1995

ABSTRACT **Agarose-encapsulated, metabolically active, permeabilized nuclei from human hematopoietic cell lines were tested for Z-DNA formation in the β-globin gene cluster. Biotinylated monoclonal antibodies against Z-DNA were diffused into the nuclei and cross-linked to DNA with a 10-ns laser exposure at 266 nm. Following digestion with restriction enzymes, fragments that had formed Z-DNA were isolated. Seventeen regions with Z-DNA sequence motifs in the 73-kb region were studied by PCR amplification, and five were found in the Z conformation.**

DNA in the chromatin of eukaryotic cells is not in a static conformation. Biological activities, such as replication and transcription as well as nucleosome formation result in the production of torsional strain, an energy source which can result in conformational change. One of the more dramatic of these changes is the conversion of right-handed B-DNA to the left-handed Z-DNA conformation (1, 2), a process that is stabilized by negative supercoiling. It has been possible to demonstrate the formation of Z-DNA in living systems, as well as in those that are very close to that found in the living cell (3–5). Using metabolically active permeabilized nuclei (6), we measured Z-DNA in an individual gene, *MYC*, and showed that its transcription is associated with the formation of Z-DNA in three discrete regions (7). Downregulation of MYC transcription during cell differentiation resulted in a rapid loss of Z-DNA. Chemical modification of the isolated Z-forming regions when negatively supercoiled made it possible to determine the nucleotides involved in Z-DNA formation (8).

In the present work we have extended this to a cluster of genes located in the same topological domain: the human β-globin gene cluster covering over 73 kb of DNA. To simplify the work, we have carried out a computer search of this region to find sequences most likely to form Z-DNA. These regions have been surveyed in a human erythropoietic cell line (K-562), as well as in a derivative cell line further differentiated along the erythropoietic path. As a nonerythropoietic cell line with previously established Z regions in the *MYC* gene, U-937 cells were analyzed. By analyzing cDNA from these cells lines, we could see that most of the early developmental forms of β-globin were expressed in K-562, but not in its adult form. The derivative cell line expressed all the β-globins of the cluster. Surveying the Z-DNA-forming sites leads to observations suggesting that the formation of Z-DNA in a topological domain with several different genes is likely to be influenced by interactions among the various genes. This type of interaction might be anticipated because the amount of Z-DNA formation is dependent on the level of transcription (9). Furthermore, the positive and negative supercoiling induced in adjacent transcribing genes will strongly influence Z-DNA formation (10).

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.

MATERIALS AND METHODS

Cell Culture, Nuclear Preparation, and Z-DNA Sequence Isolation. Hematopoietic cell line K-562 [American Type Culture Collection (ATCC); CCL 243], K-562/ADM60 and myelomonocytic U-937 cells (ATCC; CRL 1593) were grown in 90% (vol/vol) RPMI medium 1640/10% (vol/vol) fetal bovine serum (Life Technologies, Grand Island, NY) supplemented with antibiotics. K-562/ADM60 cells were generated by selecting against adriamycin at 60 ng/ml, resulting in a multidrug-resistant phenotype. They were determined to be further differentiated in erythropoietic cell maturation that K-562 cells.

The method used to isolate DNA fragments which contained Z-DNA segments from permeabilized nuclei under physiological conditions was identical to published procedures (7). PCR was carried out in 50 μl of 50 mM KCl/1.5 mM $MgCl_2$/20 mM Tris·HCl (pH 8.4)/0.8 mM dNTPs (200 μM each)/4 μM PCR primers (2 μM each) with 1.25 units of *Taq* DNA polymerase (Life Technologies) and 1 μl of DNA template solution at 0.1 μg/μl. Genomic DNA was isolated by standard procedures (11). Cycles of amplification were carried out as follows: 92°C for 1 min, probe-specific annealing temperature for 50 s, and 72°C for 20 s. Detectable bands with genomic DNA were produced after 30 cycles; 24 cycles were sufficient with a Z-fraction template. Aliquots (8 μl) were analyzed by agarose gel electrophoresis and ethidium bromide staining.

A number of DNA primers 17–25 bp in length were designed (12). The primers are described by their positions in the β-globin gene cluster (GenBank accession no. J00179). The region and primer number in parentheses are followed by the nucleotide numbers. Z1, (39) 11233–11252 and (40) 11428–11448; Z2, (1) 13143–13162 and (2) 13375–13394; Z3, (23) 26640–26659 and (24) 25648–25671; Z4, (25) 31805–31824 and (26) 31219–31238; Z5, (57) 35596–35611 and (58) 35744–35759; Z6, (41) 38378–38397 and (42) 38582–38601; Z7, (59) 40530–40549 and (60) 40632–40653; Z9, (43) 46574–46594 and (44) 46854–46871; Z10, (29) 48293–48312 and (30) 47999–48018; Z11, (45) 51820–51842 and (46) 52145–52163; Z12, (47) 52897–52919 and (48) 55180–55199; Z13, (31) 57591–57610 and (32) 56832–56851; Z14, (49) 58184–58204 and (50) 58657–58676; Z15, (51) 58842–58863 and (52) 59084–59109; Z16, (33) 59564–59583 and (34) 59169–59188; Z17, (35) 61753–61771 and (36) 61090–61109; Z18, (37) 72913–72932 and (38) 72250–72271; I, (61) 9751–97570 and (62) 9992–10011; II, (63) 26726–26745 and (64) 26842–26861; and III, (67) 59623–59642 and (68) 59717–59736. The annealing temperature was 57°C for all primers except as follows: Z3, 55°C; Z6, Z9, Z15, and Z17, 52°C; and Z10, 40°C. The primer pairs 61 and 62, 63 and 64, and 67 and 68 were controls that did not contain Z-motif sequences. A *MYC* primer pair was used as a Z-DNA control with positions (myc11) 846–861 and (myc101) 605–624 (GenBank accession no. X00364).

Abbreviation: RT/PCR, reverse transcription–PCR.

cDNA Synthesis and Expression of β-Type Globin Genes. Total cellular RNA was prepared (11). Reverse transcription (RT) was performed at 37°C for 1 h in 50 μl of 37.5 mM KCl/1.5 mM $MgCl_2$/25 mM Tris·HCl (pH 8.3)/10 mM dNTPs (2.5 mM each)/1 μM random hexamers (Pharmacia) as primers/10 mM dithiothreitol/5 μg of total cellular RNA with 200 units of Superscript reverse transcriptase (Life Technologies). The reaction was terminated by incubation at 95°C for 10 min, and the product was stored at −20°C. The annealing temperature of the PCR primer pairs and expected product length (in parentheses) are listed for the genes. The listed positions are genomic, while the product length comes from the cDNA. ε-globin: 19539–19559 and 21049–21072, 55°C (567 bp); γ-globin: 34533–34549 and 36036–36057, 55°C (517 bp); δ-globin: 54791–54808 and 56369–56382, 50°C (567 bp); β-globin: 62186–62205 and 63720–63740, 55°C (575 bp); *Myc* 7027–7046 and 5032–5051, 62°C (639 bp). One globin primer is located in the first exon; the other is in the third exon. The *MYC* primers are located in the second and third exons. A total of 35 PCR cycles were applied: 92°C for 1 min, annealing temperature for 1 min, and 72°C for 1 min. Aliquots (8 μl) of PCR products were analyzed as above. PCR products from pooled reaction mixtures were purified (QIAquick PCR Purification Kit; Qiagen) and analyzed with restriction enzymes.

RESULTS

Globin genes are expressed exclusively in cells that differentiate into erythrocytes. Their expression is regulated depending on the state of differentiation (13, 14). The globin gene cluster represents a model system in which genes are switched on during differentiation in contrast to the *MYC* gene, the expression of which is downregulated in differentiation (7). The human β-globin gene cluster spans 73 kb in one topological domain. The cluster contains five genes and one pseudogene encoded on the same DNA strand. Those genes closest to the 5′ end of the cluster are expressed earliest in development; those at its 3′ end are expressed last. The genes are arranged in the following order: embryonic (ε), fetal (Gγ, Aγ), and adult (δ, β) (Fig. 1*A*).

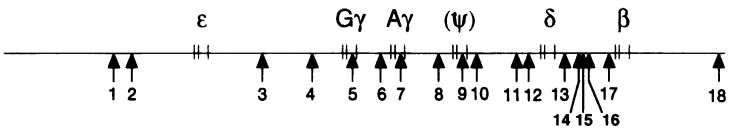
Z-DNA in the β-globin gene cluster was investigated in erythropoietic cells K-562 and K-562/ADM60. K-562/ADM60 have a cell population enriched for further erythropoietic maturation. The myelomonocytic cell U-937 served as a nonerythropoietic control. We studied β-type gene expression and searched for Z-forming DNA segments. Fig. 1*B* shows a table of computer-predicted Z segments within the β-globin gene cluster. The search algorithm is a simple heuristic calculation (12). A minimal score of 13 was selected to catch the most probable Z sequences. The program yielded 18 segments favoring the formation of Z-DNA, 17 of which were studied (Fig. 1*A*).

Fig. 2*A* shows the RT/PCR experiments carried out to measure expression of β-type globin genes. For the γ-globin genes, the primer pair used is specific for Gγ as well as for Aγ. U-937 cells showed no β-type globin gene expression, while *MYC* gene expression was detectable in all cell lines. Embryonic and fetal β-globin genes were expressed in both erythropoietic cell lines with an intensity similar to that of the *MYC* gene. Amplification of δ-globin cDNA results in a weak signal for K-562 cells and a broad signal for K-562/ADM60 cells. β-globin cDNA was detectable with weak signal intensity only in the more differentiated K-562/ADM60 cells. To verify the expression data in view of the β-globin similarities, a panel of restriction enzymes was used on the PCR products. Fig. 2B shows the estimated fragment lengths after restriction with the respective enzyme, and all except the smallest fragments were detected. For γ-globin, it can be seen that both Gγ and Aγ were amplified with the primer pair globin13/14. It appears that Gγ is more strongly represented.

PCR primer pairs for 17 *Alu* I restriction fragments containing predicted Z sites (Z1–Z7 and Z9–Z18) were selected. A primer pair specific for the first Z-forming segment of the *MYC* gene was used as a reference (7). One control primer pair (globin61/62) was specific for an *Alu* I restriction fragment arbitrarily chosen from the human β-globin gene cluster but not neighboring a fragment containing a Z-forming segment. Another pair, globin63/64, is specific for the *Alu* I fragment located next to Z3, and primer pair globin67/68 is located within the *Alu* I fragment that is two removed from segment Z16.

A subset of PCR experiments is shown in Fig. 3. The left row of agarose gels shows control experiments with genomic U-937 or K-562 DNA as a PCR template to demonstrate primer efficiency. The expected product lengths were obtained in all cases. Agarose gels presented in the right row represent

A



B

Z-segment	position	length	larg. seg.	score	Z-segment	position	length	larg. seg.	score
Z1	11303	16	14	13	Z10	48064	27	27	16
Z2	13220	37	37	24	Z11	52137	15	8	15
Z3	26355	18	18	16	Z12	53385	18	18	14
Z4	31472	22	22	18	Z13	57028	17	17	16
Z5	35557	33	33	38	Z14	58531	15	15	14
Z6	38466	15	15	13	Z15	59030	14	12	13
Z7	40493	29	27	26	Z16	59475	37	33	34
Z8	44362	14	7	13	Z17	61575	52	52	18
Z9	46665	27	24	13	Z18	72757	18	18	16

FIG. 1. Graphical representation of the human β-globin genes and prediction of Z-DNA forming sites by an algorithm. (*A*) The human β-globin gene cluster (73,326 bp) is shown with its five β-type genes (ε, Gγ, Aγ, δ, β) and one pseudogene (ψ). Each gene consists of three exons, the positions of which are depicted by vertical lines. The arrows mark the positions of 18 putative Z segments. (*B*) A table of computer-predicted Z segments within the β-globin gene cluster is shown. Prediction was performed by using the ZSEARCH algorithm (12). The algorithm calculates scores by searching a sequence of alternating purine/pyrimidine residues (minimal length: 6 bp) and then adding values for each alternating doublet. Parameters were set as follows: nonalternating doublet = −1; AT, RY, or NN = 0; AC or GT = 1; GC = 2. In the table, predicted Z segments with a minimal score of 13 are shown and named Z1–Z18. The column headed larg. seg. refers to the longest uninterrupted stretch of alternating purines and pyrimidines.

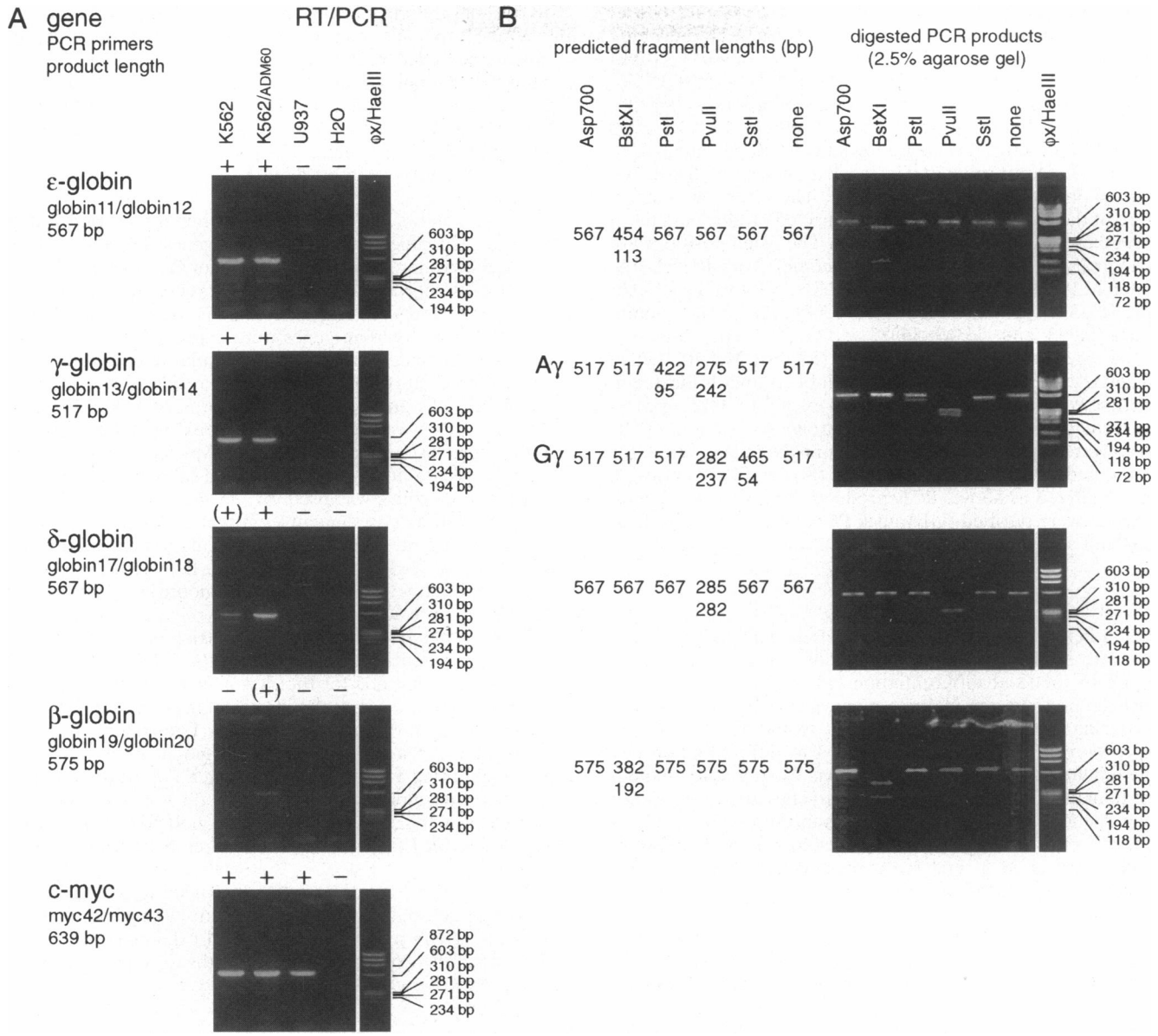


FIG. 2. Expression of genes of the β-globin cluster in the human cell lines investigated. (*A*) Amplification of β-globin mRNAs by reverse transcription–polymerase chain reaction (RT/PCR). On the left, the β-globin types, the type-specific primer pairs for PCR, and expected lengths of the amplification products are listed. *MYC* (c-myc) expression is used as a reference in view of known Z-DNA-forming sites in the *MYC* gene. The right side shows ethidium bromide-stained 2% agarose gels after electrophoretic separation of RT/PCR products. Cell lines used as mRNA sources are denoted above lanes. "H_2O" refers to RT/PCR experiments carried out in the absence of cDNA. Lane labeled φx/HaeIII contains *Hae* III-digested fragments of phage φX174 replicative form DNA of which the fragment lengths in base pairs are shown. Signs (+ or −) above each lane correspond to signal intensity. A weak β-globin signal is found in the adriamycin-treated cells. (*B*) Verification of expression data by restriction analysis of DNA produced by RT/PCR of mRNA. The left panel lists fragment lengths expected upon digestion with the respective restriction enzyme shown at the top. The right panel shows electrophoretic separation on 2.5% agarose gels. Restriction enzymes used are denoted above lanes. "none" refers to experiments carried out in the absence of restriction enzyme. Size markers as in *A*.

experiments in which DNA bound to Z-DNA-specific antibodies was used as a PCR template. A detectable signal corresponds to a DNA segment that had been in the Z conformation within the respective cell line, and was crosslinked to the antibody by the laser. The experiments presented in this figure illustrate four different types of results: (*i*) In none of the cell lines is a PCR product detectable for the predicted Z segment, as shown for Z2. Similar results (not shown) were obtained for Z1, Z5–Z7, Z9, and Z11–Z15. Likewise, no signals were detectable in control experiments using primer pairs globin61/62 and globin67/68 (results not shown). (*ii*) Segment Z16 is detectable in the Z-DNA fraction of both erythropoietic cell lines but not in U-937. Comparable results were obtained for Z3, Z4, or Z10 (results not shown). However, the signal obtained with Z16 from K-562 is much weaker than the signal obtained with K-562/ADM60. PCR products obtained for Z3, Z4, or Z10 have similar intensities in both erythropoietic cell lines (results not shown). (*iii*) The predicted Z17 segment was detectable with weak intensity only in the Z-DNA fraction of K-562/ADM60. (*iv*) PCR products are detectable in all cells. This result was obtained for Z18, for the *MYC* gene, and in a control experiment with primer pair globin63/64 (results not shown).

DISCUSSION

It is known that all the genes of the human β-globin domain are transcribed in the same direction, and they are believed to

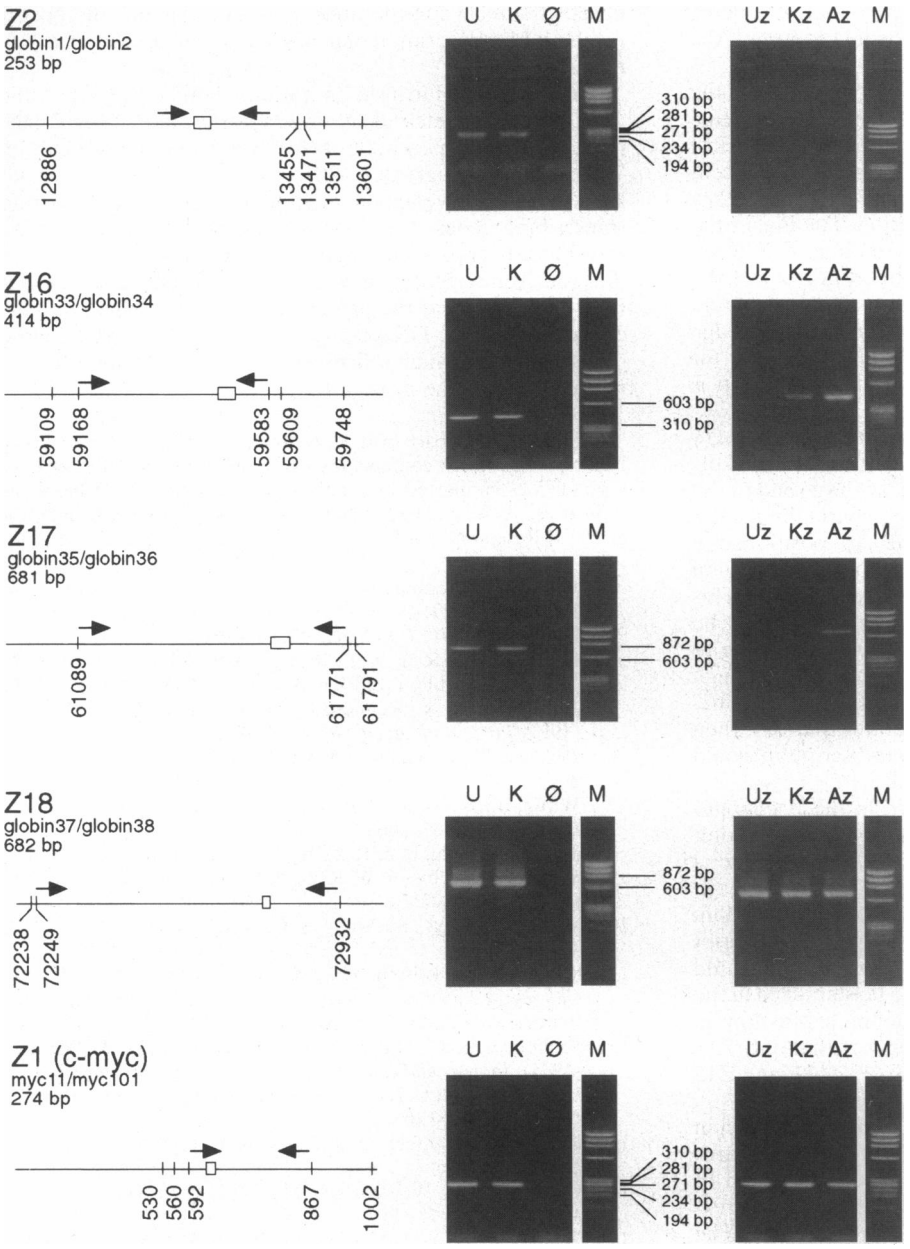


FIG. 3. Detection by PCR of Z-DNA-forming segments within the β-globin gene cluster. Schematic drawings of the sequence regions surrounding the Z sites investigated are shown on the left (Z8 was not analyzed). The predicted Z sites (numbers refer to Fig. 1), primer pairs for PCR, and expected lengths of the amplification products are listed. Horizontal arrows represent PCR primers, rectangles indicate the Z-forming DNA segment, and *Alu* I restriction sites are shown by vertical strokes (numbers correspond to positions within the β-globin gene cluster and the *MYC* gene, respectively). Z-forming DNA segment Z1 of the *MYC* gene is used as a reference. Ethidium bromide-stained agarose gels after electrophoretic separation of PCR products are shown on the right. Lane labeled M contains *Hae* III-digested phage φX174 replicative form DNA. The fragment lengths are shown in base pairs. The other lanes are labeled with letters corresponding to the DNA used as template in the amplification reactions: genomic DNA from U-937 (U) and K-562 (K), water (Ø), and isolated Z fractions from U-937 (Uz), K-562 (Kz) and K-562/ADM60 (Az).

be in one topological domain (13, 14). The organization roughly parallels their utilization during development, starting with an embryologically expressed gene and progressing to genes expressed in the fetus and finally in the adult. The erythropoietic K-562 cells weakly express the embryonic and fetal globins and only one adult gene. However, its adriamycin-resistant derivative expresses both adult β-globin forms although at different levels. In contrast, U-937 cells express *MYC* at a high level, and three regions were demonstrated that form Z-DNA while the gene is transcribed (7). Once the gene is downregulated, the Z-DNA segments disappeared rapidly. However, no comparisons were made with adjacent genes. Here, we are looking at an ensemble of five gene products separated over 50 kb. Sites with a high potential for forming Z-DNA were selected by using a computer algorithm over a 73-kb segment of DNA extending beyond the five genes. The technique for Z-DNA identification involved the laser crosslinking of biotin-labeled Z-DNA-specific monoclonal an-

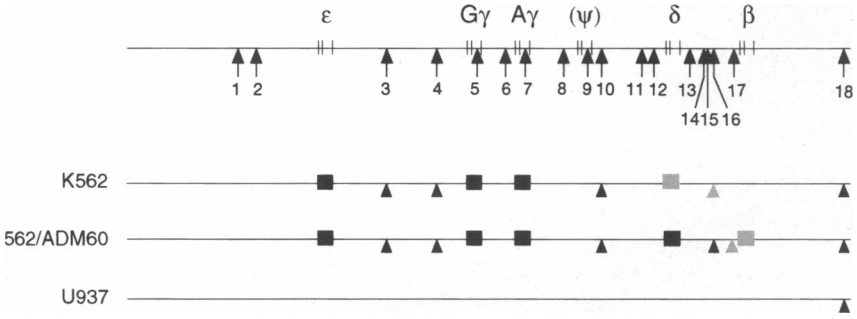


FIG. 4. Graphical representation of results. Human β-globin gene cluster is shown at the top as in Fig. 1. Results of Z-segment mapping within the β-globin gene cluster and expression of β-globin type genes are shown underneath for each cell line investigated. Expression of a single gene is depicted by rectangles. Triangles represent mapped Z segments. Black symbols indicate strong PCR signals; grey symbols correspond to weak signals.

tibodies to the DNA while the genes were actively being transcribed (7). *Alu* I digestion made it possible to isolate the segments forming Z-DNA (Fig. 4).

The major mechanism for generating Z-DNA is due to the negative supercoiling behind a moving polymerase molecule (9, 10). For genes in a linked topological domain, we must consider the effects of simultaneous transcription of adjacent genes and the effect that they would have on each other. A key issue in this regard is the rate of transcription. The higher the level of RNA synthesis, the larger the amount of Z-DNA formed (9). Positive supercoiling is generated in front of the moving polymerase, and this will impact on the torsional strain near the next downstream gene (10). However, our knowledge of these interactions is partial at best. We do not know, for example, whether the torsional strain is localized before it is distributed along the genome (15, 16).

One possible interpretation of these results is that the Gγ gene is being transcribed at a high rate. This would be unlikely to form Z-DNA in segment Z5, which is near the 3′ end of the gene; however, it could generate negative supercoiling stabilizing Z-DNA in segments Z3 and Z4. They are positioned in a region which is similar in distance to the 5′ upstream region of Z-DNA that forms when *MYC* is transcribed (7). This suggests that the transcriptional rate for ε-globin may be modest. It may not generate enough positive supercoiling to relax segments Z3 and Z4, nor does it generate enough negative supercoiling to keep Z1 and Z2 in the Z conformation. If the transcriptional rate of Gγ is greater than Aγ, then the positive supercoiling generated by Gγ transcription would prevent segment Z6 from being stabilized.

We see that Z16 and Z17 are stabilized in the Z configuration (less for Z17) associated with weak expression of adult β-globin in the K-562/ADM60 cells. In K-562 itself, no signal is generated by β-globin gene expression. Z17 is no longer in the Z conformation, and Z16 is only weakly so. The δ-globin transcript is known to be only 2–3% of the β-globin transcript in adults. If this reflects a reduced transcription rate, it would correlate with the failure of Z11 and Z12 to be stabilized in the Z conformation. The low level of β-globin expression in K-562/ADM60 may be correlated with the fact that only Z16 and Z17 are in the Z conformation, while Z13, Z14, and Z15 remain in the B conformation.

Why is the Z10 segment stabilized in the Z conformation in both erythropoietic cell lines? A reasonable suggestion regarding the Z-DNA signal at Z10 is that it is near an origin of replication, generating negative superhelicity (9). Further experiments would be needed to confirm such a surmise. The Z18 signal, positive in all cells, is located in an *Alu* repeat. We cannot rule out contamination from the ubiquitous *Alu* repeats. A similar consideration probably explains the positive results with the primer pair 63/64, which occurs in a KPN repeat.

This study has intrinsic limitations. Unlike the *MYC* study (7), the entire stretch of DNA was not examined for Z-DNA formation. An experiment of that type would be difficult in a 73-kb region. A greater difficulty stems from the limited knowledge we have about the interaction of widely separated transcribing genes. We do not know how torsional strain is transmitted over such large distances and recognize that different transcription rates could profoundly alter these interpretations. What the present study does reveal is an interesting pattern of Z-DNA formation over the extended β-globin gene cluster, although a firm interpretation of why it forms is not available at the present time.

This material is part of a Ph.D. thesis (V.M., Freie Universität, Berlin, 1995), which contains the experimental data not shown. The research was supported by grants from the Fonds der Chemischen Industrie to B.W. and by the National Institutes of Health and Office of Naval Research to A.R.

1. Rich, A., Nordheim, A. & Wang, A. H.-J. (1984) *Annu. Rev. Biochem.* **53,** 791–846.
2. Hill, R. J. (1991) *J. Cell Sci.* **99,** 675–680.
3. Rahmouni, A. R. & Wells, R. D. (1989) *Science* **246,** 358–363.
4. Wittig, B., Dorbic, T. & Rich, A. (1989) *J. Cell Biol.* **108,** 755–764.
5. Jimenez-Ruiz, A., Requena, J. M., Lopez, M. C. & Alonso, C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 31–35.
6. Jackson, D. A., Yuan, J. & Cook, P. R. (1988) *J. Cell Sci.* **90,** 365–378.
7. Wittig, B., Wölfl, S., Dorbic, T., Vahrson, W. & Rich, A. (1992) *EMBO J.* **11,** 4653–4663.
8. Wölfl, S., Wittig, B. & Rich, A., *Biochim. Biophsy. Acta*, in press.
9. Wittig, B., Dorbic, T. & Rich, A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2259–2263.
10. Wu, H.-Y., Shyy, S., Wang, J. C. & Liu, L. T. (1988) *Cell* **53,** 433–440.
11. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed., pp. 9.14–9.23.
12. Schoeneberg, U., Vahrson, W., Priedemuth, U. & Wittig, B. (1994) *Analysis and Interpretation of DNA and Protein Sequences Using MacMolly® Tetra* (Karoi, Reiner Bornemann, Bielefeld, Germany), 3rd Ed.
13. Evans, T., Felsenfeld, G. & Reitman, M. (1990) *Annu. Rev. Cell Biol.* **6,** 95–124.
14. Collins, F. S. & Weissman, S. M. (1984) *Prog. Nucleic Acid Res. Mol. Biol.* **31,** 315–462.
15. Tsao, Y.-P., Wu, H.-Y. & Liu, L. F. (1989) *Cell* **56,** 111–118.
16. Dröge, P. & Nordheim, A. (1991) *Nucleic Acids Res.* **19,** 2941–2946.