Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 5291–5295, September 1984
Biochemistry

Distinctive sequence organization and functional programming of an *Alu* repeat promoter

(superhelicity/5′ flanking region/S1 nuclease mapping/enhancing element/directing element)

CARLOS PEREZ-STABLE, THOMAS MERRILL AYRES, AND CHE-KUN JAMES SHEN*

Department of Genetics, University of California, Davis, CA 95616

Communicated by P. K. Stumpf, April 4, 1984

ABSTRACT **Plasmid clones containing a human *Alu* family repeat can be transcribed efficiently by RNA polymerase III in HeLa cell extract. This generated three RNA species, all of which initiated from the first base (+1) of the repeat. By studying the transcriptional properties of deletion clones, subclones, and topologically different DNA templates, we demonstrated that: (*i*) supercoiled DNA templates are transcribed 3- to 5-fold more efficiently than are linear or nicked circular DNA molecules; (*ii*) a contiguous DNA helix in the transcription complexes that extends into the 5′ flanking region of positions −30 to −85 is absolutely required for initiation to occur (this interaction does not involve recognition of specific DNA sequences); and (*iii*) similar to the adenovirus VAI RNA and tRNA genes, the *Alu* repeat 3′ to the α1-globin gene (designated 3′-α1 *Alu*) contains a split intragenic promoter: an anterior element (positions +4 to +37) and a posterior element (positions +70 to +82). However, the promoter of the *Alu* repeat functions in distinctive ways in comparison to those of other RNA polymerase III-dependent genes. The posterior promoter element alone is sufficient and necessary for an accurate initiation to occur. The presence of the anterior promoter element, which by itself does not initiate transcription, enhances the transcriptional efficiency by a factor of 10- to 20-fold. Furthermore, the distance between the initiation sites and the posterior promoter element, but not the anterior promoter element, remains constant. These results suggest that the promoter of this *Alu* family repeat consists of at least two functionally different domains: a "directing element" (the posterior promoter element) that determines the accuracy of initiation and an "enhancing element" (the anterior promoter element) that is mainly responsible for the transcriptional efficiency.**

In eukaryotic cells, RNA polymerase III is responsible for the transcription of genes coding for 5S rRNA, tRNA, and several small viral RNAs, including the adenovirus VAI RNA (for references, see refs. 1–3). The biological roles of 5S rRNA and tRNA in protein synthesis are well documented, whereas the VAI RNA is required for the efficient translation of late viral mRNA (4).

The development of oocyte injection techniques (5) and cell-free transcription systems (6–8) has greatly aided the identification of promoter sequences and protein factors involved in the expression of these RNA polymerase III-dependent genes (class III genes). One novel feature of the transcriptional control regions of class III genes is their location within the coding sequences (intragenic promoter). For example, a contiguous DNA block (from position +50 to +83) is essential for accurate and efficient transcription of *Xenopus* 5S rRNA genes (9, 10). A 37-kDa protein factor has been purified and shown to bind to the intragenic promoter, a process necessary for the expression of the 5S rRNA gene (11–13). Similar to 5S rRNA gene, the promoter for transcription of VAI RNA gene is also mapped within the gene at positions +10 to +72 (14, 15). Interestingly, the 5′ and 3′ portions of the VAI RNA gene promoter are homologous to the boxes A and B of the split promoters of tRNA genes, respectively (see below).

The tRNA genes have been studied in more detail (16–33). Their promoters are split into two regions located within the coding sequences (21–23). The anterior promoter region contains a highly conserved sequence 5′ T-G-G-C-N-N-A-G-T-N-G-G 3′ (box A) located 11–18 base pairs (bp) downstream from the 5′ end of the gene. The posterior promoter region also contains a conserved sequence 5′ G-G-T-T-C-G-A-N-N-C-C 3′ (box B), which is located 33–42 bp downstream from the box A sequence (21, 28). Transcriptional studies of substitution mutants of tRNA genes (17, 27, 30, 32) have confirmed the essentiality of these two boxes and their immediately flanking sequences in transcription. The spacing between the two boxes can vary without a significant loss of the transcriptional efficiency (19, 21, 22, 24, 29). Furthermore, transcription of these "insertion mutants" with varied lengths between boxes A and B always initiates upstream from the box A at positions similar to those of the wild type.

The *Alu* sequences, a major family of repetitive DNA in the human genome (34, 35), are transcribed *in vitro* by RNA polymerase III in cell-free extract (36–42). They have an approximate repetition frequency of 300,000 copies dispersed throughout the human genome and may play a role in genomic rearrangement and evolution (43, 44). A model accounting for their highly dispersed arrangement hypothesizes the expression of *Alu* family members by RNA polymerase III and subsequent synthesis of cDNA copies by reverse transcriptase (45, 46). These cDNA copies are then dispersed throughout the genome by insertion. Thus, a detailed study of the *Alu* family repeat promoter sequence would provide essential information on the molecular basis of class III transcription. It also will give us insight of an evolutionary process responsible for the genomic dispersion of a eukaryotic repetitive sequence.

MATERIALS AND METHODS

Construction of BAL-31 Deletion Mutants. The 5′ deletion mutants (see Fig. 2) were constructed by *Bcl* I cleavage of pRBα1 (grown and isolated from GM48, a *dam*⁻ strain) followed by BAL-31 nuclease (47) digestion to different extents. The 3′ deletion mutants were generated by cutting pRBα1 with *Eco*RI and then digesting the linearized molecules with BAL-31. Both sets of BAL-31-digested DNA samples were treated with Klenow fragment of polymerase I, recircularized with T4 DNA ligase (Bethesda Research Laboratories), and used to transform *Escherichia coli* strain RRI. The locations of deletions in these plasmid DNAs were determined by chemical sequencing (48).

Construction of Subclones. The plasmid pBS-Alu (−85/+62)

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.

Abbreviations: bp, base pair(s); nt, nucleotide(s).

*To whom all correspondence should be addressed.

contains a 147-bp *Bcl* I–*Sau*3A fragment (see Fig. 2) with the *Sau*3A site converted to a blunt end with Klenow enzyme. The fragment was inserted into the *Bam*HI and blunt-end *Eco*RI sites of pBR322 vector. The same vector was used in the cloning of plasmid pBT-Alu (−85/+82), which contains a 167-bp *Bcl* I–*Taq* I fragment (see Fig. 2) with the *Taq* I site converted to a blunt end with Klenow enzyme. The plasmid pHR-Alu (−30/+373) (Fig. 2) was constructed by cloning the 403-bp *Hin*cII–*Eco*RI fragment (see Fig. 1*A*) into the *Pvu* II and *Eco*RI sites of pBR322.

Synthesis and Analysis of *in Vitro* Transcripts. Cytoplasmic S-100 transcription extracts were prepared and transcription reactions were performed as described by Wu (8). RNAs were analyzed by electrophoresis on 7 M urea/6% polyacrylamide gel and autoradiography.

S1 Nuclease Mapping of 5′ Ends of *in Vitro* Transcripts. DNA fragments were 5′-end-labeled with [α-^{32}P]ATP (Amersham) by using T4 polynucleotide kinase (Biotec, Madison, WI). The appropriate single-stranded DNA probes were isolated by elution from 7 M urea/8% polyacrylamide gels. S1 nuclease mapping experiments were performed as described (49–51).

RESULTS

A plasmid subclone (pRBα1, Fig. 1*A*) contains an *Alu* family member located on the 3′ side of human α1-globin gene, designated 3′-α1 *Alu*. *In vitro* in HeLa S-100 extract, the *Alu* repeat is transcribed at least 20-fold more efficiently than are most of the other *Alu* family members (41). The generation of the three transcripts a, b, and c (Fig. 1*B*, lane 2) has been shown by S1 nuclease mapping and primer extension techniques to be the result of unique initiation at the first base of the repeat and multiple termination downstream (Fig. 1*A*; unpublished data).

Supercoiled DNA Templates Are Transcribed More Efficiently. The plasmid pRBα1 has a single *Bcl* I restriction site located at position −85 upstream from the *Alu* transcriptional initiation site and a single *Eco*RI site at 373 bp downstream (Fig. 1*A*). Linearization of pRBα1 at the *Bcl* I restriction site results in a transcription pattern similar to that of pRBα1 but with reduced band intensities (20–30%) of all three RNA species (Fig. 1*B*, lane 4). Similar observations were made with pRBα1 linearized by *Eco*RI except that the RNA-a species was substituted by a shorter run-off transcript [approximate length, 370 nucleotides (nt)] (Fig. 1*B*, lane 3). We also prepared nicked pRBα1 DNA, and the transcriptional pattern and band intensities were similar to the *Bcl* I-linearized pRBα1 (unpublished observation). The same phenomenon also was observed for a subclone pHR-Alu (−30/+373) (compare lanes 7 and 8 in Fig. 1*B*). These results demonstrate that the presence of supercoil turns enhances the transcriptional efficiencies *in vitro* of the *Alu* repeat by a factor of 3–5. This magnitude of difference is similar to the *in vivo* value of tRNA genes (52) but is considerably less than RNA polymerase II-dependent genes (53).

Transcriptional Initiation Requires a Contiguous DNA Helix in the 5′ Flanking Region from −30 to −85. When pRBα1 is cleaved with the restriction enzyme *Hin*cII, the initiation of transcription is totally abolished (Fig. 1*B*, lane 5). This is not caused by inhibiting contaminants in the preparation of *Hin*cII-cleaved pRBα1 DNA because a sample of *Eco*RI-linearized pRBα1 can still direct accurate initiation in the presence of *Hin*cII-cleaved pRBα1 (Fig. 1*B*, lane 6) as judged from the presence of the 370-nt RNA but not the 410-nt-long RNA-a.

We have subcloned the 403-bp *Hin*cII–*Eco*RI fragment (Fig. 1*A*) into *Pvu* II and *Eco*RI sites, respectively, of pBR322. This has brought in a totally different DNA sequence upstream from position −30. It was transcribed accurately and efficiently in the HeLa S-100 extract (Fig. 1*B*, lane 7). Thus, a contiguous nucleotide sequence extending into the region between −30 and −85 is essential for the transcriptional initiation to occur. However, this interaction does not involve recognition of specific DNA sequences. This is consistent with previous observations on VAI gene and several tRNA genes by different experimental approaches (15, 18, 20, 25, 31). The 5′ flanking sequences of those genes seem to have a modulating effect on transcription.

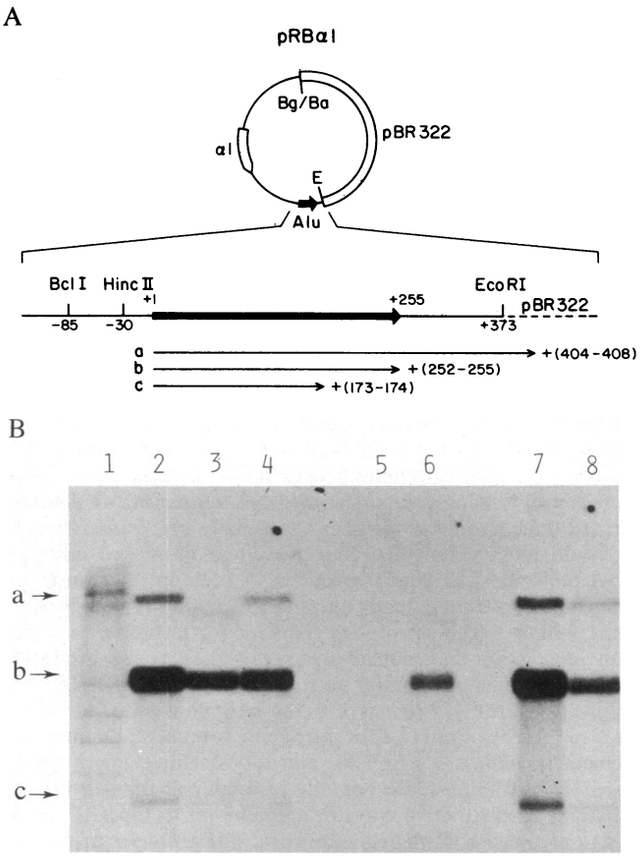

FIG. 1. (*A*) Transcription map of the *Alu* family repeat located on the 3′ side of α1-globin gene. The plasmid pRBα1 contains a 4-kilobase-long, *Bgl* II (Bg)–*Eco*RI (E) human DNA fragment cloned into the *Bam*HI (Ba) and *Eco*RI (E) sites of pBR322. This DNA insert contains the human α1-globin gene (the blank arrow) and an *Alu* family repeat (designated 3′-α1 Alu, the black arrow) located downstream from the globin gene. *In vitro* transcription of pRBα1 results in the synthesis of RNAs a, b, and c as indicated by the three thin arrows below the linear restriction map. The numbers in the parentheses to the right of the arrows indicate the lengths (in nt) of each RNA species. A minus sign indicates a position is upstream from the transcriptional initiation site and a plus sign indicates a downstream position. (*B*) *In vitro* transcription of 3′-α1 *Alu* repeat. Different DNA samples (2 μg each) containing the 3′-α1 *Alu* repeat were transcribed in HeLa S-100 extract, and the RNA products were analyzed by gel electrophoresis and autoradiography. Lanes: 1, single-stranded ^{32}P-labeled *Hae* III-cut pBR322 fragments as size markers; 2, pRBα1; 3, *Eco*RI-cut pRBα1; 4, *Bcl* I-cut pRBα1; 5, *Hin*cII-cut pRBα1; 6, *Hin*cII-cut pRBα1 and *Eco*RI-cut pRBα1; 7, pHR-Alu (−30/+373); 8, *Pst* I-cut pHR-Alu (−30/+373).

Intragenic Promoter of 3′-α1 *Alu* Repeat Contains at Least Two Functionally Different Domains. Previous sequence comparison (28, 37, 39, 41) and transcriptional study of subcloned restriction fragments (40) have suggested the existence of an intragenic promoter of *Alu* family repeats. The promoter of the 3′-α1 *Alu* repeat has been mapped in detail by *in vitro* transcription of BAL-31 deletion clones or subclones of pRBα1 constructed as described. The maps of these DNA clones and their transcriptional behavior in S-100

extract are summarized in Fig. 2.

An *in vitro* transcription gel of the 3′ deletion clones is shown in Fig. 3*A*. Deletion of the original termination sites and the use of new termination sites (clusters of thymidine residues on the sense strand) in pBR322 sequence brought in after ligation have resulted in the new RNA patterns. The sizes of the major RNA species are consistent with the notion that initiation in these clones occurs at the same position as in pRBα1. Deletions beyond base +140 [e.g., clones p3d (+63/+551), p3d (+47/+483), and p3d (+18/+558)] abolish the transcription completely. These data, as summarized in Fig. 2, indicate that the 3′ border of the promoter is located in between positions +63 and +140 of the *Alu* sequence. However, a subclone pBT-Alu (−85/+82) can be transcribed efficiently (Fig. 3*C*). Thus, the 3′ border of the promoter is located in between positions +63 and +82. Consistent with these results, the subclone pBS-Alu (−85/+62) is not transcribed at all in the extract (Fig. 3*C*).

Transcription of the 5′ deletion clones is shown in Fig. 3*B*. Deletion of all the 5′ flanking sequences and 3 nt past the normal initiation site has no effect on transcription. Deletions extending into positions +37, +50, and +69 [in clones p5d (−248/+37), p5d (−334/+50), and p5d (−219/+69)] all resulted in weak transcription with the efficiencies being ≈5–10% of the wild type. However, the transcriptional patterns of these three 5′ deletion clones are all similar to pRBα1, indicating their transcriptional initiation sites are located at approximately the same positions as those of the wild type, even though different DNA sequences had been brought in by the deletion and cloning treatment. Transcription was eliminated upon deletion to position +79 [see clone p5d (−296/+79) in Fig. 3*B*].

To ascertain the accuracy of initiation of the three weakly transcribed 5′ deletion clones, p5d (−248/+37), p5d (−334/+50), and p5d (−219/+69), we have used the S1 nu-

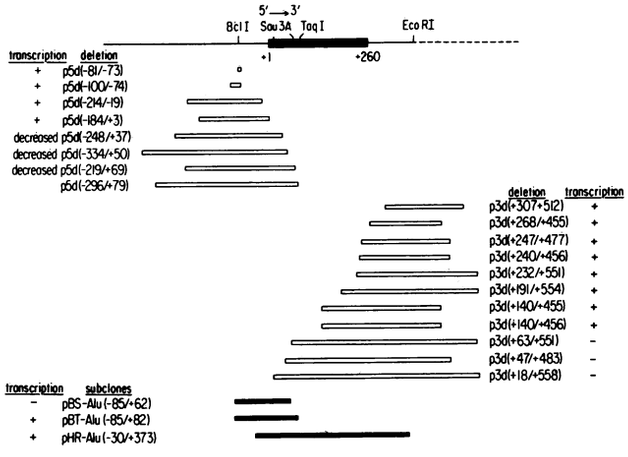

FIG. 2. Summary of transcriptional properties of different BAL-31 deletion clones and subclones constructed from pRBα1. The 3′-α1 *Alu* repeat is represented at the top by the black bar. The thin solid line represents human DNA sequence flanking the repeat, while the dashed line is pBR322 DNA sequence. *BAL*-31 deletion clones were constructed from *Bcl* I-linearized pRBα1 (p5d series) and from *Eco*RI-linearized pRBα1 (p3d series). The regions of DNA sequences removed by *BAL*-31 nuclease in each deletion clone is represented by the open bars. The numbers in parentheses refer to the boundaries of deleted DNA sequences relative to the first base (+1) of 3′-α1 *Alu*, which is also the transcriptional initiation site of the repeat. DNA inserts contained within the three subclones are represented by the hatched bars. The transcriptional properties of these clones in HeLa S-100 extracts are indicated: "+", accurate and efficient initiation; "decreased," an accurate but inefficient initiation; and "−", no transcriptional initiation. These results were derived from the data shown in Figs. 3 and 4.

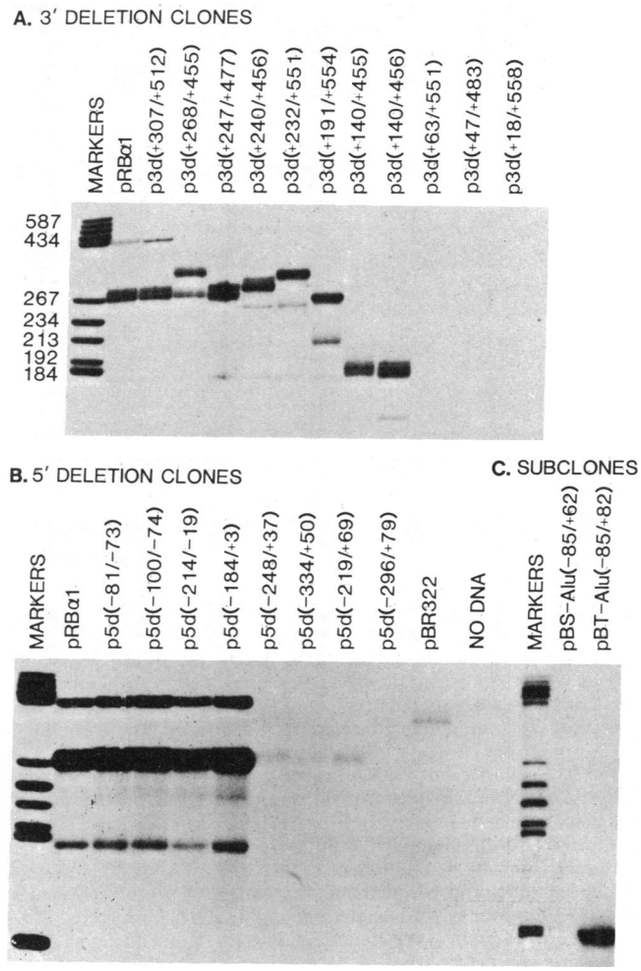

FIG. 3. (*A*) *In vitro* transcription products of the 3′ deletion clones (p3d series in Fig. 2) analyzed on 7 M urea/6% polyacrylamide gel. Markers are single-stranded, ^{32}P-end-labeled *Hae* III-cut pBR322 fragments with the sizes (in nt) indicated to the left. Numbers in parentheses are explained in the legend to Fig. 2. (*B*). *In vitro* transcription products of the 5′ deletion clones (p5d series in Fig. 2). A longer exposure time was necessary to clearly see the weak transcription patterns of p5d (−248/+37), p5d (−334/+50), and p5d (−219/+69) when compared to pRBα1. (*C*). *In vitro* transcription of the two subclones pBS-Alu (−85/+62) and pBT-Alu (−85/+82). The 120-nt RNA from pBT-Alu (−85/+82) is the result of initiation at the first base of the *Alu* repeat (see text and Figs. 4 and 5) and termination in a cluster of thymidine residues in the pBR322 vector sequence.

clease-digestion technique to map the 5′ end(s) of their major transcripts. The strategy of this mapping experiment and the autoradiographs are shown in Fig. 4 *A* and *B*, respectively. All three 5′ deletion clones and pBT-Alu (−85/+82) initiate at approximately the same positions as does pRBα1 (Fig. 5). The S1-mapping results have recently been confirmed by primer extension technique (unpublished results).

From all these data, we conclude that the promoter of 3′-α1 *Alu* repeat is located within the coding sequence and consists of at least two functionally different domains (Fig. 5). One of them, which we term "directing element," is located between bases +70 and +82. The presence of this directing element is sufficient and necessary for programming an accurate initiation, with the initiation sites being ≈70 bases upstream from the 5′ border of the element. The second domain is contained in between bases +4 and +37. Since this sequence by itself cannot initiate transcription but will greatly enhance the efficiency of transcription directed by

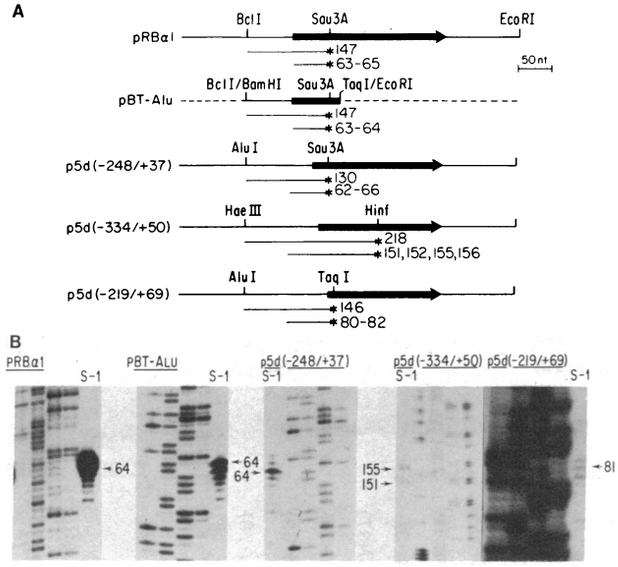

FIG. 4. (*A*) S1 nuclease mapping strategy. The 5′ ends of the major transcripts from pRBα1 and pBT-Alu (−85/+82) (both have the wild-type promoter) and the three weakly transcribed 5′ deletion clones were mapped by the S1 nuclease-protection technique. The thick arrow represents the 3′-α1 *Alu* DNA sequence, while the thin line represents the flanking human DNA sequence. The dashed line in the map of pBT-Alu (−85/+82) represents the sequence from pBR322. Single-stranded DNA fragments used to hybridize with the RNA samples and their lengths (in nt) are shown below the individual map, with an asterisk indicating the labeled 5′ ends. The sizes of protected DNA fragments after nuclease S1 digestion have been determined from the gel patterns of Fig. 4*B* and are shown below each of the full-length probes. Numbers in parentheses are explained in the legend to Fig. 2. (*B*). Size determination of RNA-protected DNA probes on 7 M urea/8% polyacrylamide gel. DNA sequencing markers were coelectrophoresed to measure the sizes of protected probes. In each panel, the sequencing reactions (from left to right) are G, A+G, C+T, and C. The lengths (in nt) of the protected DNA fragments are indicated by numbers on the side of each panel. The initiation sites of these five clones as determined from this experiment are shown schematically in Fig. 5.

the "directing element," we termed it "enhancing element."

DISCUSSION

In summary, we have studied the effect of either DNA conformation or sequence arrangement on the transcriptional accuracy and efficiency of a human *Alu* family repeat. The most surprising result from this study is the distinctive functional organization of *Alu* repeat promoter.

Transcriptional analysis of deletion clones and subcloned restriction fragments (Figs. 2, 3, and 4) revealed the existence of two essential promoter elements located within the *Alu* repeat sequence (Fig. 5). One of these elements, the "enhancing element," is located between +4 and +37 and contains within it a sequence (G-G-C-G-C-G-G-T-G-G) homologous to the conserved box A region of tRNA gene promoter. The other promoter sequence, the "directing element," is located downstream from the enhancing element at positions +70 to +82. This element, G-T-C-A-G-G-G-G-T-T-C-G-A, overlaps for 6 bp with a sequence homologous to the conserved box B of tRNA promoter (see Fig. 5). Thus, the general sequence organization of the *Alu* repeat promoter is very similar, although not identical, to the other class III genes, especially to the split promoter of tRNA genes. However, the programming of transcriptional initiation by the *Alu* intragenic promoter elements is distinct from that of the tRNA genes in two major aspects. First, the posterior promoter element (the directing element) alone is sufficient and necessary for an accurate initiation to occur. The presence of the anterior element (the enhancing element) increases the transcriptional efficiency by a factor of 10- to 20-fold. Second, the distance between the initiation site(s) and the posterior element, but not the anterior element, remains the same regardless of the upstream DNA sequences. These properties are certainly different from the way tRNA gene promoter functions. The transcription of many eukaryotic tRNA genes and the VAI gene requires a cooperative interaction of both the box A- and box B-containing promoter sequences (14–16, 19, 21–24, 28, 33). In homologous systems, the presence of either the anterior or posterior control region alone does not give a positive transcriptional initiation. Furthermore, in tRNA genes, the initiation sites remain approximately constant distances upstream from box A instead of box B (19, 21, 22, 24, 29).

One primary interaction required for tRNA gene transcription is binding of transcription factor(s) to the box B region and its immediate flanking sequences. Crude protein fractions from yeast cells have been demonstrated by DNase I-protection analysis to interact specifically with 20–30 bp of DNA centered around the box B sequence of yeast tRNA genes and VAI gene (32, 54, 55). Given the high homology of the box B region of the promoters among these genes, it is very likely that they share common cofactor(s) essential for transcription. This has been shown to be the case with both competition binding assay and transcriptional analysis (15, 22, 33, 54, 55). Removal of the box A region of the tRNA promoters results in the loss of template activity. However, these clones are still able to compete in the transcription extract with wild-type genes for limiting factors essential for transcription. Removal of the box B region of the promoter results in the loss of competition for these factors. Single-point mutations in the first half of the box B sequence of tRNA and *Alu* repeat promoter regions have been shown to eliminate transcription (32, 42), possibly because of the lack of binding of a specific cofactor. Thus, it is apparent that the box B-containing region of VAI gene, tRNA genes, and *Alu* repeats is the "primary" sequence element of the promoter.

Considering their similar sequence organization, it is not

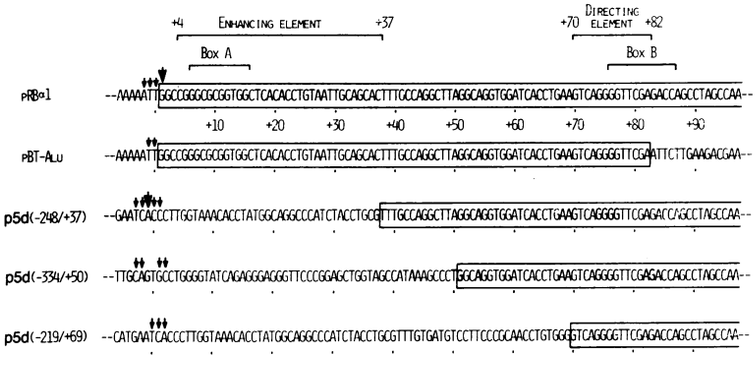

FIG. 5. Summary map showing the locations of the two functionally different promoter domains, the enhancing element, and the directing element of 3′-α1 *Alu* repeat. Also shown are positions of initiation sites (arrows above each DNA sequence) as determined by the S1 nuclease-digestion experiment of Fig. 4. The boxed regions contain the *Alu* repeat sequence. For pRBα1, the larger arrow indicates the initiation site as determined by primer extension technique (unpublished results). The box A and box B sequences indicated are homologous to the two-well-conserved sequences of tRNA gene and VAI RNA gene promoters (14, 21, 27).

obvious at this moment why the promoter sequences of the tRNA genes and the *Alu* family repeat function in different ways. We do notice, however, that the 13-bp-long posterior directing element of the 3′-α1 *Alu* repeat contains 7 bp of the 5′ portion of the box B-like sequence and an additional 6-bp sequence upstream (Fig. 5). This 6-bp sequence is homologous but not identical to the consensus of tRNA gene sequences immediately upstream from box B. It is possible that the nucleotide sequence difference between *Alu* repeats and other class III genes within this DNA segment is responsible for their difference in directing the transcription *in vitro*. This is consistent with the study by Allison *et al.* (27). They have shown that six of the seven single-base substitutions surrounding the posterior promoter regions in tRNA gene mutants having altered transcriptional activities or competition abilities for transcriptional factors are clustered within the 5′ portion of the box B consensus sequence and its immediate upstream sequences. A change of the nucleotide sequence of this segment of DNA in most tRNA genes or VAI gene could have abolished the ability of their posterior promoter elements to independently direct transcriptional initiation. In these genes the positioning of the initiation sites is then played by the anterior promoter element in cooperation with the posterior element. Finally, it is interesting to point out a similarity in the organization between promoters of *Alu* family repeats and RNA polymerase II-dependent genes. Although the promoters of the RNA polymerase II-dependent genes are located in the 5′ flanking region, they also consist of two functionally different elements. The downstream element mainly controls the accuracy of initiation, while the upstream element is responsible for the transcriptional efficiency (for example, see refs. 56 and 57).

We thank John Hess for his help in the S1 nuclease mapping and primer extension experiments. We also appreciate Peter Yao and Morton Bradbury for letting us use their cell culture facilities at the beginning stage of this project. This research has been supported by a grant from The National Institutes of Health (AM 29800). C.P.-S. is a National Institute of Health Biochemistry Trainee (5T32-GM07377-06-AI).

1. Sklar, V. E. F., Yamamoto, M. & Roeder, R. G. (1976) in *RNA Polymerases*, eds. Losick, R. & Chamberlin, M., (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Mongr. 6, pp. 803–817.
2. Jaehning, J. A., Weinmann, R., Brendler, T. G., Raskas, H. T. & Roeder, R. G. (1976) in *RNA Polymerases*, eds. Losick, R. & Chamberlin, M., (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Monogr. 6, pp. 819–834.
3. Ciliberto, G., Castagnoli, L. & Cortese, R. (1983) *Curr. Top. Dev. Biol.* **18,** 59–87.
4. Thimmappaya, B., Weinberger, C., Schneider, R. J. & Shenk, T. (1982) *Cell* **31,** 543–551.
5. Brown, D. D. & Gurdon, J. B. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2849–2853.
6. Birkenmeier, E. H., Brown, D. D. & Jordan, E. (1978) *Cell* **15,** 1077–1086.
7. Weil, P. A., Segall, J., Harris, B., Ng, S. Y. & Roeder, R. G. (1979) *J. Biol. Chem.* **264,** 6163–6173.
8. Wu, G. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2175–2179.
9. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. (1980) *Cell* **19,** 13–25.
10. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. (1980) *Cell* **19,** 27–35.
11. Engelke, D. R., Ng, S. F., Shastry, B. S. & Roeder, R. G. (1980) *Cell* **19,** 717–728.
12. Pelham, H. R. & Brown, D. D. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4170–4174.
13. Honda, B. M. & Roeder, R. G. (1980) *Cell* **22,** 119–126.
14. Fowlkes, D. M. & Shenk, T. (1980) *Cell* **22,** 405–413.
15. Guilfoyle, R. & Weinmann, R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3378–3382.
16. Kressman, A., Hofstetter, H., DiCapua, E., Grosschedl, R. & Birnstiel, M. L. (1979) *Nucleic Acids Res.* **7,** 1749–1762.
17. Koski, R. A., Clarkson, S. G., Kurjan, J., Hall, B. D. & Smith, M. (1980) *Cell* **22,** 415–425.
18. Sprague, K. U., Larson, D. & Morton, D. (1980) *Cell* **22,** 171–178.
19. Corrara, G., DiSegni, G., Otsuka, A. & Tocchini-Valentini, G. P. (1981) *Cell* **27,** 371–379.
20. DeFranco, D., Sharp, S. & Söll, D. (1981) *J. Biol. Chem.* **256,** 12424–12429.
21. Galli, G., Hofstetter, H. & Birnstiel, M. L. (1981) *Nature (London)* **294,** 626–631.
22. Hofstetter, H., Kressman, A. & Birnstiel, M. L. (1981) *Cell* **24,** 573–585.
23. Sharp, S., DeFranco, D., Dingermann, T., Farrell, P. & Söll, D. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6657–6661.
24. Ciliberto, G., Traboni, C. & Cortese, R. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1921–1925.
25. Dingermann, T., Burke, D. J., Sharp, S., Schaack, J. & Söll, D. (1982) *J. Biol. Chem.* **257,** 14738–14744.
26. Hall, B. D., Clarkson, S. G. & Tocchini-Valentini, G. (1982) *Cell* **29,** 3–5.
27. Allison, D. S., Goh, S. H. & Hall, B. D. (1983) *Cell* **34,** 655–665.
28. Ciliberto, G., Raugei, G., Costanzo, F., Dente, L. & Cortese, R. (1983) *Cell* **32,** 725–733.
29. Dingermann, T., Sharp, S., Shaack, J. & Söll, D. (1983) *J. Biol. Chem.* **258,** 10394–10402.
30. Folk, W. R. & Hoffstetter, H. (1983) *Cell* **33,** 585–593.
31. Hipskind, R. A. & Clarkson, S. G. (1983) *Cell* **34,** 881–890.
32. Newman, A. J., Odgen, R. C. & Abelson, J. (1983) *Cell* **35,** 117–125.
33. Sharp, S., Dingermann, T., Schaack, J., DeFranco, D. & Söll, D. (1983) *J. Biol. Chem.* **258,** 2440–2446.
34. Schmid, C. W. & Jelinek, W. (1982) *Science* **216,** 1065–1070.
35. Jelinek, W. & Schmid, C. W. (1982) *Annu. Rev. Biochem.* **51,** 813–844.
36. Duncan, C. H., Biro, P. A., Choudary, P. V., Elder, J. T., Wang, R. R. C., Forget, B. G., deRiel, J. K. & Weissman, S. M. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 5095–5099.
37. Pan, J., Elder, J. T., Duncan, C. H. & Weissman, S. M. (1981) *Nucleic Acids Res.* **9,** 1151–1170.
38. Fritsch, E. F., Shen, C.-K. J., Lawn, R. M. & Maniatis, T. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45,** 761–765.
39. Duncan, C. H., Jagadeeswaran, P., Wang, R. R. C. & Weissman, S. M. (1981) *Gene* **13,** 185–196.
40. Fuhrman, S. A., Deininger, P. L., LaPorte, P., Friedman, T. & Geiduschek, E. P. (1981) *Nucleic Acids Res.* **9,** 6439–6456.
41. Shen, C.-K. J. & Maniatis, T. (1982) *J. Mol. Appl. Genet.* **1,** 343–360.
42. Murphy, M. H. & Baralle, F. E. (1983) *Nucleic Acids Res.* **11,** 7695–7700.
43. Hess, J. F., Fox, M., Schmid, C. W. & Shen, C.-K. J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5970–5974.
44. Schmid, C. W. & Shen, C.-K. J. (1984) in *Molecular Evolutionary Genetics*, ed. McIntyre, R. J. (Plenum, New York), in press.
45. Van Arsdell, S. W., Denison, R. A., Bernstein, L. B., Weiner, A. M., Manser, T. & Gesteland, R. F. (1981) *Cell* **26,** 11–17.
46. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. (1981) *Cell* **26,** 141–142.
47. Legerski, R., Hodnett, J. & Gray, H. B., Jr. (1978) *Nucleic Acids Res.* **5,** 1445–1464.
48. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
49. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12,** 721–732.
50. Weaver, R. F. & Weissman, C. (1979) *Nucleic Acids Res.* **7,** 1175–1193.
51. Favaloro, J. M., Treiman, R. H. & Kamen, R. (1980) *Methods Enzymol.* **65,** 718–749.
52. Melton, D. (1978) Dissertation (Cambridge Univ., Cambridge).
53. Harland, R. M., Weintraub, H. & McKnight, S. L. (1983) *Nature (London)* **302,** 38–43.
54. Klemeny, R., Stillman, D. & Geiduschek, P. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6191–6195.
55. Lassar, A. B., Martin, P. L. & Roeder, R. G. (1983) *Science* **222,** 740–748.
56. Struhl, K. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7385–7389.
57. McKnight, S. L. & Kingsbury, R. (1982) *Science* **217,** 316–324.