Volume 11 Number 22 1983 Nucleic Acids Research

Directed semisynthetic point mutational analysis of an RNA polymerase III promoter

M.H.Murphy and F.E.Baralle

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received 6 October 1983; Accepted 31 October 1983

ABSTRACT

The transcription of tRNA and Alu repeat genes in vitro by RNA polymerase III has been shown to be dependent on the presence of two intragenic regions, which contain the consensus sequences RGYNNRRYGG (box A) and G$^{A}_{T}$TCRANNC (box B), located 30-60 nucleotides apart. The role of box B and some of its variants was analysed by a novel method involving the chemical synthesis of double stranded analogues of box B which were subsequently cloned into recombinant vectors carrying box A alone. This method creates a series of semi-synthetic RNA polymerase III promoters and has no limitation on the structure and number of variants which can be generated. The results showed the "wild type" sequence GTTCGAGAC and the sequence GTTCGTGAC (an A → T transversion of the 6th position) were active in promoting RNA polIII transcription. However, the box B sequences CTTCGAGAC and GTACGAGA, where the only departures from the consensus are a G → C and an A → T transversion in the 1st and 3rd positions respectively, were unable to restore promoter function.

INTRODUCTION

The genes transcribed by RNA polymerase III contain their promoter within the coding sequence. A map of the regions essential for promotion is available for the 5S and tRNA genes (1-6). We have previously established that the efficient transcription of the RNA polymerase III promoter of the intermediate repetitive DNA Alu family is dependent upon two regions containing the consensus sequences RGYNNRRYGG (box A) and G$^{A}_{T}$TCRANNC (box B) located approximately 60 nucleotides apart (2). We concluded that the Alu RNA polymerase III promoter has a bipartite structure which resembles tRNA. The boxes were first noted by comparison of sequence homologies between all functional RNA polymerase III promoters, while the boundaries of regions essential for transcription were defined by experiments relying either on the availability of suitable restriction sites or random mutations (1,2,6-8). These approaches have made slow progress in identifying short sequences or nucleotides which are essential for transcription. For example, only one of the nucleotides found in box B has been directly demonstrated in vitro to be essential for

transcription (9). To overcome these experimental limitations, we chose to test directly the role of the box B consensus sequence in transcription by cloning a chemically synthesized mixture of duplex box B-related oligonucleotides into a subclone of the Alu promoter which lacked the naturally occurring box B.

MATERIALS AND METHODS

Oligonucleotide Synthesis

The mixture of double stranded box B-related oligonucleotides was synthesized as the complementary strands 5' GGAGTTCGAGACCAC 3' (C A T) and 5' GTGGTCTCGAACT (A T G) CC 3' by the solid phase phosphotriester method (10). We included the 3 bases flanking either side of the box B consensus sequence to allow for a highly selective hybridisation screen (11).

Clone Construction

The M13 recombinants 1 and 2 (see Figure 1) have been described (2). Clones 3-10 were constructed by cloning the chemically synthesized box B duplexes

```
          C A T
5' - GGAGTTCGAGACCAC - 3'
3' - CCTCAAGCTCTGGTG - 5'
       G T A
```

into the filled-in EcoR1 site of clone 2. The duplexes were unphosphorylated to avoid the formation of concatamers. Typically a hundred fold molar excess of oligo duplex over vector was employed, for example, a 10μl ligation consisted of 50ng vector, 10ng oligo, 5 units T4 ligase, 1mM rATP, 50mM Tris-HCl pH 7.4, 10mM $MgCl_2$, 12mM DTT. The correct orientation of box A with respect to box B was selected for by a hybridisation between the phage plaques and the complementary oligonucleotide (11). The positive clones were sequenced by the dideoxy method (12).

In vitro transcription of the clones and analysis of the products

The source of RNA polymerase III used was an S100 supernatant from HeLa cells prepared according to Weil et al (13). All the DNAs tested as template were titrated up to 3μg of DNA in a 25μl reaction mixture containing 15μl of S100 supernatant, 0.5mM ATP, CTP and UTP, 0.05mM GTP and 5μCi of [α-^{32}P] GTP (Amersham International sp. act. 410Ci/mM), 1mM creatine phosphate and 0.1mM EDTA. The reaction mixtures were incubated for 60 min. at 30°C and then processed according to Birkenmeier et al (14). The RNA products were then fractionated by electrophoresis on a 6% acrylamide 7M urea gel.

RESULTS AND DISCUSSION

The subclone used as a vector contains only the 5' half of the promoter (box A and flanking sequences, see Fig. 1 line 2), and was previously demonstrated to be inactive as an RNA polymerase III template (see ref. 2 and Fig. 2). The mixture of double stranded box B-related oligonucleotides was cloned in the "filled-in" EcoRl site of clone 2 (Fig. 1), thus mimicking the naturally occurring relationship between box A and B. Seven out of the possible 8 box B-related sequences were isolated (see Figure 1); however, the 3' G of the oligonucleotide GTTCTGTGAC consensus box was always deleted in this experiment (clone 4, Figure 1). In addition, a wild type consensus box B (GTTCGAGAC) was isolated where the 3' AG of the oligonucleotide was deleted (clone 10, Fig. 1).

These DNAs were tested as RNA polymerase III templates in the HeLa S100

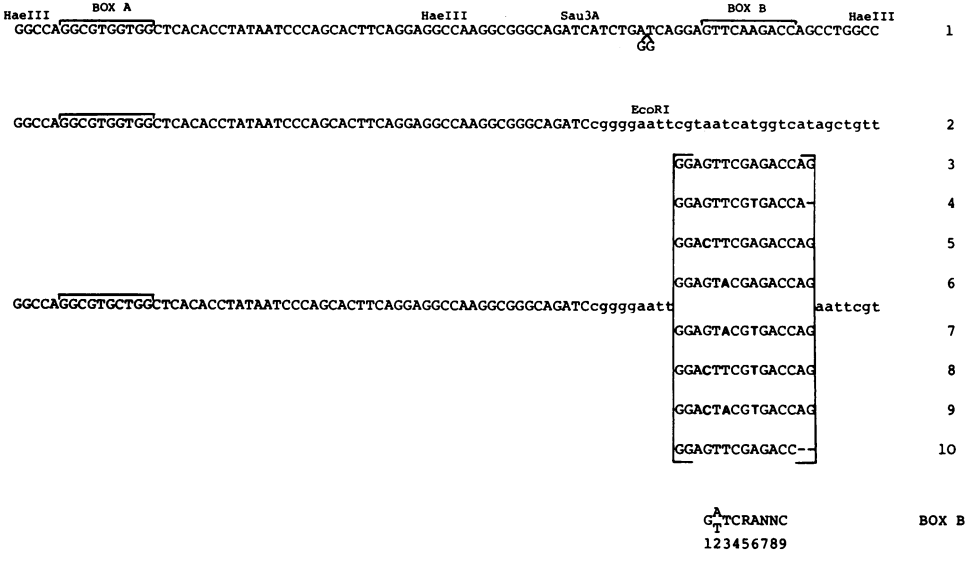

Figure 1 Nucleotide sequence of the Alu repeat promoter and the variations constructed.

The DNA sequences corresponding to the RNA transcript are shown. The molecules read 5' to 3' from the left of the figure. Clone 1 is the fully active natural Alu repeat promoter; boxes A and B are indicated. Clone 2 is a derivative of clone 1 where box B is deleted; the M13 vector sequence is in lower case type. Clones 3-10 were constructed as described in Materials and Methods. The box B sequence permutations of each recombinant isolated are indicated within the brackets. Variations from the wild type sequence are in bold type. Small dashes have been introduced to facilitate alignment. The relevant restriction enzyme sites are indicated.

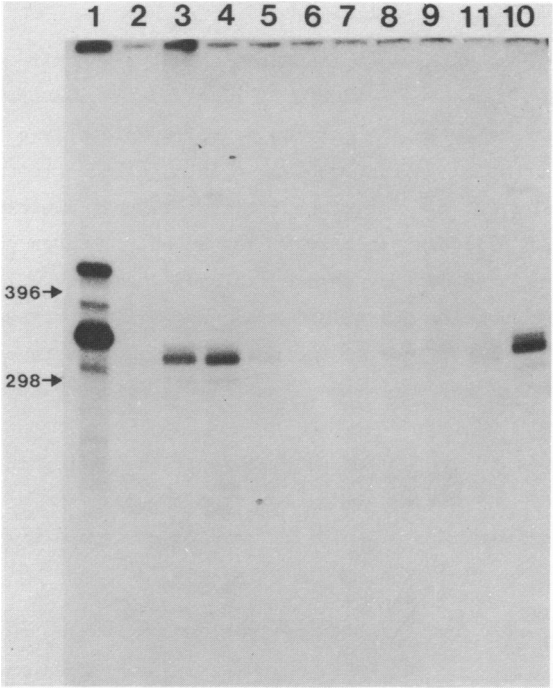

Figure 2 Autoradiograph of a 6% 7M urea gel fractionation of RNA polymerase III transcripts from clones 1-11 (see Figure 1).

Single stranded size markers are indicated. The major transcript observed was approximately 330 nucleotides long, consistent with transcription initiation 5-15 nucleotides 5' to box A as occurs in the natural Alu repeat, and termination at a strong stop in the M13 vector sequence.

in vitro system described by Weil et al (Figure 2)(13). Lane 1 shows the transcripts produced by the parent Alu repeat cloned in pBR322 (see reference 2), while lanes 2-10 show transcripts produced by the individual recombinants represented in Figure 1 lines 2-10. Template activity was not detected either when box B was absent (clone 2 lane 2, Figure 2) or when box B wild type sequence was present alone in the M13 vector (lane 11, Fig. 2). Promoter activity was restored in clone 3 where the chemically synthesized wild type box B was introduced into clone 2 (Fig. 2, lane 3). A similar activity was detected in clone 10 (Fig. 2, lane 10) where a wild type consensus sequence lacking the 3' AG of the synthetic oligonucleotide is present. The major transcript observed was approximately 330 nucleotides long, consistent with transcription initiation 5-15 nucleotides 5' to box A as occurs in the natural Alu repeat, and termination at a strong stop in the M13 vector

sequence (see reference 2). This was confirmed using the in vitro produced RNA as template and the mixture of oligonucleotides complementary to box B as primer for cDNA synthesis with reverse transcriptase (15). A band 90 nucleotides long was observed which places initiation in the predicted region (data not shown).

An A → T transversion of the 6th position of the consensus sequence (clone 4, Fig. 2) similarly allowed promoter function. The additional nucleotide deletion at the 3' end of the oligonucleotide in this clone is unlikely to be significant as it lies outside the consensus box; the equivalent activities of clones 2 and 10, which differ by a similar deletion, further indicate that a deletion at the 3' end of the oligonucleotide does not affect transcription.

It is clear that a single G → C or T → A transversion at positions 1 and 3 respectively (clones 5 and 6) significantly reduces transcription (see Fig. 2 lanes 5 and 6). In the case of multiple mutations, where two or three single mutations were introduced into the consensus B box (clones 7-9), no template activity was observed (see Fig. 2, lanes 7-9). This indicates that the G → C or the T → A transversions are dominant and abolish the promoter function regardless of any other change in the consensus sequence (like the A → T transversion).

We have been able to study variations of the RNA polymerase III promoter box B using a new approach to introduce mutations in a consensus sequence. The results directly confirm for the first time the central role of the wild type box B consensus sequence in transcription. Furthermore two of the invariant nucleotides of the consensus sequence, G_1 and T_3, were shown to be essential for promoter activity. It was also shown that there is no absolute requirement of A_6 for efficient transcription, although this nucleotide is totally conserved in all functional tRNA and Alu promoters. This contradiction could occur if this nucleotide has a functional role in the gene product or its processing; it has previously been demonstrated that a G → T transversion at position 5 of the tRNA box B consensus allows efficient transcription but only partial processing (16).

It is generally assumed that the A and B boxes interact with one or more transcription factors in addition to the polymerase. The specific recognition signals could be simply a function of the linear sequence of the DNA or could result from a specific DNA conformation (for a review, see 17). Our demonstration that the box B consensus sequence is not as stringently defined as previously proposed (1-3) weakens the argument for an exclusive primary

sequence requirement, and we propose that the consensus block B sequence should now read G$\overset{A\ \ TA}{TTCRTNNC}$. The construction of every possible variation in the A and B boxes using the same methodology will help to test more critically the available models.

ACKNOWLEDGEMENTS

We thank Dr B Barrell of the MRC Laboratory of Molecular Biology, Cambridge, for assistance on the DNA sequencing techniques, and Dr A J Caton for critical reading of the manuscript. M.H.M. is the holder of an MRC research studentship, and this work was supported by MRC grant no. G8004432CB to F.E.B.

REFERENCES

1. Hofstetter, H., Kressman, A. and Birnstiel, M.L. (1981) Cell 24, 573-585.
2. Paolella, G., Lucero, M.A., Murphy, M.H. and Baralle, F.E. (1983) EMBO Jnl. 2, 691-696.
3. Ciliberto, G., Castagnoli, L. and Cortese, R. (1983) Curr.Topics Dev.Biol. 18, 59-88.
4. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) Cell 19, 27-35.
5. Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980) Cell 19, 13-25.
6. Ciliberto, G., Raugei, G., Costanzo, F., Dente, L. and Cortese, R. (1983) Cell 32, 725-733.
7. Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) Nature 294, 626-631.
8. Ciliberto, G., Castagnoli, L., Melton, D.A. and Cortese, R. (1982) Proc. Natl.Acad.Sci.USA 79, 1195-1199.
9. Koski, R.A., Clarkson, S.G., Kurjan, J., Hall, B.D. and Smith, M. (1980) Cell 22, 415-425.
10. Gait, M.J., Singh, M., Sheppard, R.C., Edge, M.D., Greene, A.R., Heathcliffe, G.R., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1980) Nucl. Acids Res. 8, 1081-1096.
11. Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981) Nucl.Acids Res. 9, 879-894.
12. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc.Natl.Acad.Sci.USA 74, 5463-5467.
13. Weil, P.A., Luse, D.S., Segall, J. and Roeder, R.G. (1979) Cell 18, 469-484.
14. Birkenmeier, E.H., Brown, D.D. and Jordan, E. (1978) Cell 15, 1077-1086.
15. Baralle, F.E. (1977) Cell 12, 1085-1095.
16. Zasloff, M., Santos, T., and Hamer, D.H. (1982) Nature 295, 533-535.
17. Hall, B.D., Clarkson, S.G. and Tocchini Valentini, G. (1982) Cell 29, 3-5.