Volume 280, number 2, 284–286 FEBS 09513 March 1991

ADONIS 001457939100220M

Variability of polyadenylation sites in mRNAs from human fetal liver

Vladimir V. Dmitrenko and Vadim M. Kavsan

Department of Biosynthesis of Nucleic Acids, Institute of Molecular Biology and Genetics, Ukr. SSR Academy of Sciences, Zabolotnogo Street 150, 252627 Kiev, USSR

Received 17 January 1991

cDNA libraries of human fetal liver were constructed in pBR322 and λgt10 vectors. The libraries were screened for liver-specific clones by differential hybridization. This procedure revealed 25 and 32 liver-specific clones in plasmid and phage libraries, respectively. The majority of these clones were represented with serum albumin, fetal $^{G}\gamma$-globin and $^{A}\gamma$-globin cDNA inserts. Three types of 3′-non-coding region were found in 5 sequenced albumin cDNAs. In one type mRNA the distance between the AATAAA signal and polyadenylation site was 15 nucleotides, in 2 other types this distance was 10 and 6 nucleotides. The polyadenylation site in the $^{G}\gamma$-globin cDNA was located 2 nucleotides further from AATAAA signal, while in the $^{A}\gamma$-globin cDNA it was 2 nucleotides closer to the signal as compared with the results published previously.

Polyadenylation site; Differential hybridization; Albumin; γ-Globin; Liver-specific cDNAs; Human fetal liver

1. INTRODUCTION

One of mRNA maturation steps is the cleavage and polyadenylation of RNA polymerase II transcripts. This step plays an important role in the regulation of gene expression. In particular, alternative usage of multiple polyadenylation sites for the same transcription unit can lead to the generation of mRNAs of different length coding for various proteins. For instance the IgM gene produces secreted or membrane – bound antibodies [1]. As a rule alternative polyadenylation occurs by using different AATAAA signals. Mouse α-amylase [2] and dihydrofolate reductase [3] genes, human liver alcohol dehydrogenase [4] and aromatase P-450 [5] genes can serve as the examples.

Alternative polyadenylation of mRNAs under the influence of a single AATAAA signal is less investigated. The data on such alternative polyadenylation for human albumin, $^{G}\gamma$-globin and $^{A}\gamma$-globin mRNAs, discovered while sequencing of the cDNA inserts from liver-specific clones, are presented.

2. MATERIALS AND METHODS

2.1. *RNA preparation and construction of cDNA libraries*

Total RNA was prepared from human 23-week-old embryos by the guanidine thiocyanate method [6]. Poly(A) RNA fraction was isolated by passage through oligo(dT)-cellulose (Pharmacia, type 77F).

The cDNA library in pBR322 vector was constructed by the dG:dC-tailing method [7] and the cDNA library in λgt10 vector was constructed as described by Glover [8].

Correspondence address: V.M. Kavsan, Dept of Biosynthesis of Nucleic Acids, Institute of Molecular Biology and Genetics, Ukr. SSR Academy of Sciences, Zabolotnogo Street 150, 252627 Kiev, USSR

2.2. *Screening of the cDNA libraries for liver-specific clones*

Liver-specific clones were isolated by means of differential hybridization. Human fetal liver and human kidney cDNAs were used as hybridization probes. Two nitrocellulose filter sets to be hybridized were prewashed in 50% formamide, 4×SSC, 100 μg/ml denatured bovine DNA, 100 μg/ml synthetic poly(rA), and 4 × Denhardt's solution at 42°C. Hybridizations with [^{32}P]cDNA were carried out in the same buffer in sealed plastic bags at 42°C for 24 h. After hybridization the filters were washed 4 times in 2 × SSC, 0.1% SDS at 65°C and twice in 0.1 × SSC, 0.1% SDS at 65°C.

2.3. *DNA sequence analysis*

Identification of liver-specific clones was performed by dot hybridization analysis [9] and by the Maxam-Gilbert sequencing procedure [10].

3. RESULTS AND DISCUSSION

Twenty-five and thirty-two liver-specific clones were chosen from about 1000 pBR322 recombinants and from about one thousand λgt10 recombinants, respectively. The majority of these clones were represented with serum albumin, fetal $^{G}\gamma$-globin and $^{A}\gamma$-globin cDNA inserts.

The human albumin gene contains two functional polyadenylation signals. This fact was confirmed by the isolation of cDNA clones with different lengths of 3′-non-coding regions [11,12]. These signals are used differentially. The signal proximal to the protein termination codon is the dominant one [13]. We found 3 different polyadenylation sites in the sequenced cDNAs for albumin being the result of alternative utilization of proximal AATAAA signal. In clones pHA1 and pHA12 the distance between the AATAAA signal and the polyadenylation site is 15 nucleotides (as described by Dugaiczyk et al. [11]). In clones pHA8 and pHA25

Published by Elsevier Science Publishers B.V.

```
pHA1,pHA12...TTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCC
pHA8,pHA25...TTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCC
pHA19      ...TTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCC

pHA1,pHA12  TCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAATCT↓(A)n
pHA8,pHA25  TCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAG↓(A)n
pHA19       TCTTTTCTCTGTGCTTCAATTAATAAAAAATGG↓(A)n
```

Fig. 1. Nucleotide sequence comparison of albumin cDNA 3'-terminal regions from pHA1, pHA12, pHA8, pHA25, pHA19 clones. Polyadenylation signal is underlined; arrows indicate poly(A)-addition sites.

it corresponds to 10 nucleotides and in clone pHA19, 6 nucleotides (Fig. 1). The analysis of human genomic clones and DNA in previous reports [13,14] indicates that albumin is encoded by a single-copy gene. Therefore, our results most probably show that 3 revealed types of mRNA were generated as a result of the alternative processing of one primary transcript of the albumin gene.

Four clones with $^{G}\gamma$-globin and one clone with $^{A}\gamma$-globin cDNAs were sequenced. Comparison of nucleotide sequences of these cDNAs with those published previously [15–18] showed that the polyadenylation site of the $^{G}\gamma$-globin cDNAs (described herein) was located 2 nucleotides further from the AATAAA signal (Fig. 2A), whilst in the $^{A}\gamma$-globin cDNA it was 2 nucleotides closer to the signal (Fig. 2B). Two variants (I and II, Fig. 2A) of the $^{G}\gamma$-globin cDNA sequenced in our laboratory are probably the allelic cDNAs. Nucleotide sequence I is identical to the genomic nucleotide sequence described by Slighton et al. [19]. The difference between nucleotide sequences I and II is the deletion of adenosine residue located 6 nucleotides upstream of the AATAAA signal. The variant of the $^{G}\gamma$-globin gene with such a deletion was

```
A
                TER
I      ...TGA GCTCACTGCCCATGATGCAGAGCTTTCAAGGATAGGCTTTATTC
II
III                  G                  -
IV                   G                  -
V                    G                    G

I      TGCAAGCAATACAAATAATAAATCTATTCTGCTAAGAGATCAC↓(A)n
II                   -                                ↓(A)n
III                                                   ↓(A)n
IV                                                 T  ↓(A)n
V                                                     ↓(A)n

B
                TER
I      ...TGA GCCTCTTGCCCATGATTCAGAGCTTTCAAGGATAGGCTTTATTC
II     ...TGA GCCTCTTGCCCATGATTCAGAGCTTTCAAGGATAGGCTTTATTC

I      TGCAAGCAATACAAATAATAAATCTATTCTGCTGAGAGATC↓  (A)n
II     TGCAAGCAATACAAATAATAAATCTATTCTGCTGAGAGATCAC↓  (A)n
```

Fig. 2. Nucleotide sequence comparison of human $^{G}\gamma$-globin cDNA (A) and $^{A}\gamma$-globin cDNA (B) 3'-untranslated regions. Polyadenylation signal is underlined; arrows indicate poly(A)-addition sites. (A) Nucleotide sequences of $^{G}\gamma$-globin cDNAs determined in this work (I and II); by Forget [15] (III); Cavallesco et al. [16] (IV); and Lang et al. [17] (V). (B) Nucleotide sequences of $^{A}\gamma$-globin cDNAs determined in this work (I), and by Poon et al. [18] (II).

identified by Shiokawa et al. [20] when analysing the sequence heterogeneity of the γ-globin genes in Japanese individuals.

The existence of multiple cleavage/polyadenylation sites under the influence of a single signal have been reported for hepatitis B virus surface antigen [21] and for several eukaryotic genes: bovine prolactin [22]; mouse ribosomal protein L30 [23]; chicken pro-α_2(I)collagen [24]; barley toxin β-hordothionin [25], rat salivary gland α-2μ globulin [26]. Our results show the existence of multiple polyadenylation sites for several genes from the liver of one individual.

The termination of transcription in higher eukaryotes occurs beyond the normal 3'-end of a poly(A) mRNA and formation of biologically active mRNA molecules involves endonucleolytic cleavage and polyadenylation of the RNA precursors. A determination of the specific sequences and factors required for accurate 3'-end processing is the subject of investigations in many laboratories. The AATAAA hexanucleotide is one of the major elements necessary for cleavage and polyadenylation of pre-mRNAs [27]. A region extending downstream of the poly(A)-addition site is essential for formation of the 3'-end of the mRNA [28]. Many genes contain a YGTGTTYY consensus sequence in this region [29]. The observed alternative usage of the same polyadenylation signal probably reflects a complicated secondary structure of the region extending downstream of the poly(A)-addition site or the presence of additional regulatory elements in this region.

REFERENCES

[1] Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980) Cell 22, 313-319.
[2] Tosi, M., Young, R.A., Hagenbuchle, O. and Schibler, U. (1981) Nucleic Acids Res. 9, 2313-2323.
[3] Setzer, D.R., Mc Grogan, M., Nunberg, J.H. and Schimke, R.T. (1980) Cell 22, 361-370.
[4] Heden, L.O., Hoog, J.O., Larsson, K., Lake, M., Lagerholm, E., Holmgren, A., Vallee, H. and Van Bahr-Lindstrom, H. (1986) FEBS Lett. 194, 327-332.
[5] Toda, K., Terashima, M., Mitsuuchi, Y., Yamasaky, Y., Yokoyama, Y., Nojima, S., Ushiro, H., Maeda, T., Yamamoto, Y., Sagare, Y. and Shizuta, Y. (1989) FEBS Lett. 247, 371-376.
[6] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
[7] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[8] Glover, D.M. (1985) DNA Cloning Volume I, II: A Practical Approach, IRL Press, Oxford, Washington DC.
[9] Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) Basic Methods In Molecular Biology, Elsevier, New York.
[10] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
[11] Dugaiczyk, A., Law, S.W. and Dennison, O.E. (1982) Proc. Natl. Acad. Sci. USA 79, 71-75.
[12] Lawn, R.M., Adelman, J., Bock, S.C., Franke, A.E., Houck, C.M., Najarian, R.C., Seeburg, P.H. and Wion, K.L. (1981) Nucleic Acids Res. 9, 6103-6114.
[13] Urano, Y., Watanabe, K., Sakai, M. and Tamaoki, T. (1986) J. Biol. Chem. 261, 3244-3251.
[14] Hawkins, J.W. and Dugaiczyk, A. (1982) Gene 19, 55-58.
[15] Forget, B.G. (1980) Curr. Top. Hematol. 3, 1-74.
[16] Cavallesco, C., Forget, B.G. and De Riel, J.K. (1980) Gene 12, 215-221.
[17] Lang, K.M. and Spritz, R.A. (1985) Gene 33, 191-196.
[18] Poon, R., Kan, Y.W. and Boyer, H.W. (1978) Nucleic Acids Res. 5, 4625-4630.
[19] Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627-638.
[20] Shiokawa, S., Fucharoen, S., Fucharoen, G., Tomatsu, S. and Fukumaki, Y. (1989) J. Biochem. 105, 184-189.
[21] Simonsen, C.C. and Levinson, A.D. (1983) Mol. Cell. Biol. 3, 2250-2258.
[22] Sasavage, N.L., Smith, M., Gillam, S., Woychick, R.P. and Rottman, F.M. (1982) Proc. Natl. Acad. Sci. USA 79, 223-227.
[23] Wiedeman, L.M. and Perry, R.P. (1984) Mol. Cell. Biol. 4, 2518-2528.
[24] Aho, S., Tate, V. and Boedtker, H. (1983) Nucleic Acids Res. 11, 5443-5450.
[25] Hernandez-Lucas, C., Royo, J., Paz-Ares, J., Ponz, F., Garcia-Olmedo, F. and Carbonero, P. (1986) FEBS Lett. 200, 103-106.
[26] Gao, F., Endo, H. and Yamamoto, M. (1989) Nucleic Acids Res. 17 4629-4636.
[27] Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 63, 211-214.
[28] Gil, A. and Proudfoot, N.J. (1987) Cell 49, 399-406.
[29] McLauchlan, J., Caffney, D., Whitton, J.L. and Clements, J.B. (1985) Nucleic Acids Res. 13, 1347-1368.