Biochimica et Biophysica Acta 1859 (2016) 355–365

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Polyadenylation of RNA transcribed from mammalian SINEs by RNA polymerase III: Complex requirements for nucleotide sequences

Olga R. Borodulina, Julia S. Golubchikova, Ilia G. Ustyantsev, Dmitri A. Kramerov *

Laboratory of Eukaryotic Genome Evolution, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov St., Moscow 119991, Russian Federation

ARTICLE INFO

Article history:
Received 8 September 2015
Received in revised form 9 December 2015
Accepted 11 December 2015
Available online 14 December 2015

Keywords:
SINE
RNA polymerase III
Non-coding RNA
Polyadenylation
Signal nucleotide sequences
Mammals

ABSTRACT

It is generally accepted that only transcripts synthesized by RNA polymerase II (e.g., mRNA) were subject to AAUAAA-dependent polyadenylation. However, we previously showed that RNA transcribed by RNA polymerase III (pol III) from mouse B2 SINE could be polyadenylated in an AAUAAA-dependent manner. Many species of mammalian SINEs end with the pol III transcriptional terminator (TTTTT) and contain hexamers AATAAA in their A-rich tail. Such SINEs were united into Class T^+, whereas SINEs lacking the terminator and AATAAA sequences were classified as T^-. Here we studied the structural features of SINE pol III transcripts that are necessary for their polyadenylation. Eight and six SINE families from classes T^+ and T^-, respectively, were analyzed. The replacement of AATAAA with AACAAA in T^+ SINEs abolished the RNA polyadenylation. Interestingly, insertion of the polyadenylation signal (AATAAA) and pol III transcription terminator in T^- SINEs did not result in polyadenylation. The detailed analysis of three T^+ SINEs (B2, DIP, and VES) revealed areas important for the polyadenylation of their pol III transcripts: the polyadenylation signal and terminator in A-rich tail, β region positioned immediately downstream of the box B of pol III promoter, and τ region located upstream of the tail. In DIP and VES (but not in B2), the τ region is a polypyrimidine motif which is also characteristic of many other T^+ SINEs. Most likely, SINEs of different mammals acquired these structural features independently as a result of parallel evolution.

1. Introduction

It is common knowledge that all mRNAs are generated by RNA polymerase II (pol II) and subjected to polyadenylation by nuclear polyadenosine polymerase (PAP) [1]. As a result, the 3′ ends of all fully processed eukaryotic mRNAs (except for most of histone mRNAs) have a 150–250 nt poly(A) tail. Poly(A) plays an important role in regulation of stability, export from the nucleus to the cytoplasm, and translation initiation of mRNAs. Polyadenylation is associated with the cleavage of pre-mRNA resulting in the formation of mRNA 3′ end [2]. This process requires that the pre-mRNA contains (i) a polyadenylation signal (PAS), AAUAAA, 10–30 nt upstream of the cleavage site, and (ii) a U- or GU-rich element 20–40 nt downstream of the cleavage site [3]. Upstream elements (USE) located up to 100 nt of the poly(A) site are also involved in cleavage and polyadenylation of some pre-mRNAs. The cleavage and polyadenylation of pre-mRNA require at least five protein factors including PAP [4]. They are associated with the elongating polymerase complex through the C-terminal domain (CTD) of the largest subunit of pol II. Thus, these factors interact with the nascent pre-mRNA resulting in its cleavage and polyadenylation.

Until recently, only transcripts synthesized by pol II were thought to be polyadenylated in an AAUAAA-dependent manner. However, we demonstrated unambiguously that B2 RNA transcribed by RNA polymerase III (pol III) from mouse SINE B2 can be effectively polyadenylated in mammalian cells [5]. Short Interspersed Elements (SINEs), or short retroposons, are repetitive 80- to 400-bp sequences that are interspersed over the eukaryotic genomes and are amplified via reverse transcription (see reviews [6,7]). SINEs belong to nonautonomous mobile elements since they encode no enzymes and utilize the reverse transcriptase of Long Interspersed Elements (LINEs) for their amplification by mode of retroposition. Most mammalian SINEs proliferate with the help of LINE-1. The genome of mammalian species usually contains several families of SINEs, each represented by 10^4–10^6 copies. Most copies are not identical and their nucleotide sequences vary by 5–35%. A typical SINE consists of three parts: 5′-terminal "head," "body," and 3′-terminal "tail." The heads of SINEs demonstrate a similarity with one of the three types of cellular RNAs synthesized by pol III: tRNA, 7SL RNA, or 5S rRNA. Most SINE families originated from tRNA in the course of evolution. SINEs are transcribed by pol III due to the promoter in their head region. The classical SINE promoter

Abbreviations: CFI_m, Cleavage factor I_m; CPSF, Cleavage and polyadenylation specificity factor; CstF, Cleavage stimulation factor; CTD, C-terminal domain; LINE, Long Interspersed Element; PAP, Polyadenosine polymerase; PAS, Polyadenylation signal; pol II, RNA polymerase II; pol III, RNA polymerase III; PTB, Polypyrimidine tract-binding protein; SINE, Short Interspersed Element; TRAMP complex, Trf4/Air2/Mtr4p Polyadenylation complex; USE, Upstream elements.

* Corresponding author.
E-mail address: kramerov@eimb.ru (D.A. Kramerov).

http://dx.doi.org/10.1016/j.bbagrm.2015.12.003
1874-9399/

consists of two 11 bp boxes (A and B) spaced by 30–40 bp. Usually, the origin of SINE body is unknown; tails of most mammalian SINEs are A-rich sequences.

The A-rich tail of SINE B2 contains several AATAAA hexamers and a pol III transcriptional terminator (TCTTTT). This fact and finding of the short B2 transcripts in a fraction of poly(A)-containing RNA suggested that this pol III-generated RNA was polyadenylated in mouse cells [8–10]. Later, we transfected HeLa cells with B2 carrying either the normal AATAAA hexamers or the modified ones (AACAAA). It was found that the RNA transcribed by pol III from normal B2 was polyadenylated, whereas B2 RNA without AAUAAA hexamers was unable to perform this post-transcriptional modification [5]. Thus, AAUAAA can function as PAS in B2 RNA generated by pol III (Fig. 1). Polyadenylation of B2 RNA drastically increases its lifespan [5]. Additionally, a long A-rich tail (poly(A) in the case of B2 RNA) is known to be necessary for the retroposition of LINE-1-mobilized SINEs [11–13].

Like B2, thirteen mammalian SINEs have A-rich tails containing AATAAA hexamers and a pol III transcriptional terminator, $(T)_{4-6}$ or $TC(T)_{3-5}$ [14]. A-rich tails of many other mammalian SINEs lack both AATAAA and the terminator. We assigned these SINE families to class T^+ and class T^-, respectively [15].

Here we studied the polyadenylation ability of RNA transcribed from eight T^+ class SINEs. Then AATAAA hexamers and a terminator were added to tails of six T^- class SINEs. Pol III transcripts of these modified SINEs were examined for their capacity to be polyadenylated. Further, the position of PAS in B2 RNA was studied as determinant for its polyadenylation. Deletions, replacements, and translocations of various regions of B2 RNA were used to examine their significance for polyadenylation. Pol III transcripts of SINEs DIP and VES were studied in a similar way. We made the following conclusions: (i) transcripts of all T^+ class SINEs studied are effectively polyadenylated; (ii) transcripts of all T^- class SINEs studied cannot be polyadenylated even after insertion of AATAAA hexamers and a terminator in their tails; (iii) two regions in the body of SINEs B2, DIP, and VES are required for efficient polyadenylation of their transcripts; (iv) pol III transcripts capable of PAS-dependent polyadenylation have complex structural features that SINEs acquired independently in various eutherian lineages.

2. Materials and methods

2.1. Plasmid constructs

All SINEs and their derivatives containing deletions or nucleotide substitutions were obtained using PCR. Amplified DNA fragments were purified by electrophoresis in agarose gel and cloned into plasmid pGEM-T (Promega) following the manufacturer's protocol. Plasmids with cloned insertions were isolated by alkaline lysis miniprep and subjected to DNA sequence analysis in order to exclude plasmids with random nucleotide substitutions induced by PCR. Plasmids intended for transfection were isolated using NucleoBond AX100 columns (Macherey-Nagel).

2.2. Cell transfections

Transfections were carried out in the HeLa cell line. Monolayer was grown to 80% confluency in Petri dishes (60-mm diameter). Cells were transiently transfected with 5 μg of plasmid DNA using TurboFect Transfection Reagent (Thermo Scientific) following the manufacturer's protocol. RNA was isolated 20 h post-transfection using the guanidinium thiocyanate method.

2.3. RNA analysis

RNA was treated with 100 μg/mL RNase-free DNase I (Thermo Scientific) at 37 °C for 30 min. Following inactivation of DNase, RNA was resolved by denaturing electrophoresis in 5.5% polyacrylamide gel. RNA was transferred from the gel onto a Hybond-XL membrane by semidry electroblotting at 3 V for 2.5 h. Hybridization probes that usually corresponded to 5′ half of SINEs were prepared by PCR. Amplified DNA fragments were purified by electrophoresis in agarose gel and eluted with Gel Band Purification Kit (GE Healthcare). The probes were labeled with α[^{32}P] dATP by PCR using only the reverse primer. The blots were hybridized in 50% formamide, 5 × Denhardt solution, 4 × SSC, 1% SDS, and 0.1 mg/mL salmon sperm DNA at 42 °C. Washes were performed in 0.1% SSC and 0.1% SDS at 42 °C for 1 h. Then the membranes were exposed to X-ray film or/and screen (type SR) for the Phosphorimager Cyclone (Packard). The phosphorimager images were used to quantify SINE transcript polyadenylation (see Supplementary note). Radioactive signal was measured for the full-length primary SINE transcript (the band of the longest RNA) and for the polyadenylated transcript (an even longer RNA looking as a smear). Finally, the percentage of polyadenylation was calculated for the transcripts of each SINE construct. Most of the transfection experiments repeated three to eight times. Results of typical experiments are shown in figures.

3. Results

3.1. Pol III transcripts of all T^+ class SINEs can be polyadenylated

Eight unrelated SINE families from class T^+ were studied in this experiment (Table S1). A full-length SINE copy with a normal pol III promoter and a typical A-rich tail was selected from each family. Such SINE copies were cloned together with short (50–70 bp) 5′ flanking sequences (Fig. S1). From these copies, derivative constructs containing AACAAA instead of AATAAA hexamers in their A-rich tails were obtained. Pairs of the constructs were used for transient transfection of HeLa human carcinoma cells, RNAs were isolated 20 h after transfection, and SINE transcripts were detected by Northern hybridization. Change of a T with a C in all AATAAA hexamers in SINEs resulted in a quite narrow band of RNA, whereas in the case of native constructs, longer heterogeneous RNAs were also observed (Fig. 2A). These longer heterogeneous RNAs are the result of polyadenylation of the primary pol III

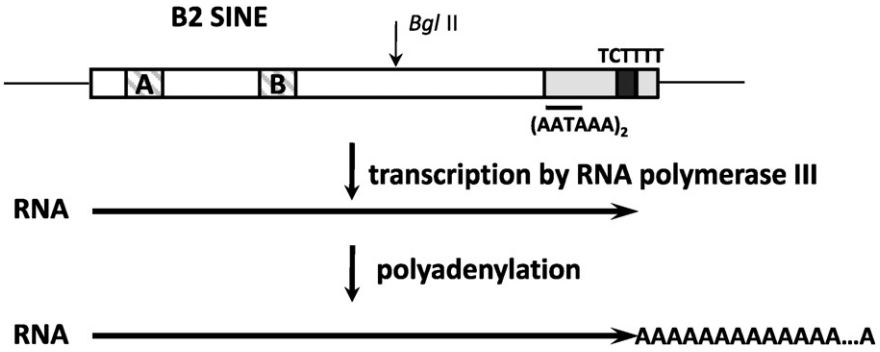

Fig. 1. Schematic of mouse B2, the best-studied SINE of class T^+. B2 transcription by pol III is initiated at the internal split promoter (boxes A and B) and is terminated at TCTTTT (solid box). AATAAA hexamers in the A-rich tail (gray box) are required for the polyadenylation of B2 RNA.