J. Indian Inst. Sci., Jan.–Feb. 1989, **69**, 65–71.
© Indian Institute of Science.

3′-Terminal polyadenylation of mRNA in procaryotes

N. SARKAR, J. TALJANIDISZ, P. KARNIK, P. SHEN AND Y. GOPALAKRISHNA

Boston Biomedical Research Institute, Department of Metabolic Regulation, 20, Staniford Street, Boston, MA 02114, USA

AND

Harvard Medical School, Department of Biological Chemistry & Molecular Pharmacology, Boston, MA 02115, USA.

Received on February 11, 1988.

Abstract

Studies in our laboratory have shown that polyadenylation of bacterial mRNA is a common phenomenon, in contrast to general belief. Poly(A)RNA constitutes a surprisingly large fraction of mRNA in *Escherichia coli*, *Bacillus subtilis*, and *Bacillus brevis*. cDNA can be synthesized from bacterial poly(A)RNA using reverse transcriptase only in the presence of oligo(dT) primer, thus providing independent evidence that bacterial mRNA is polyadenylated. This facilitated the construction of a cDNA library from poly(A)RNA of *B. subtilis*. The 3′-terminal sequences adjacent to the poly(A) moiety has been determined in some clones. We have also examined the polyadenylation of specific *E. coli* mRNAs. They include the relatively stable mRNA for the outer membrane lipoprotein (lpp) as well as the mRNA for tryptophan synthase (trpA), which serves as paradigm for relatively unstable mRNAs encoding intracellular proteins. So far, all the specific gene transcripts we have probed in *B. subtilis* and in *E. coli* were found to be polyadenylated.

Key words: Polyadenylation, bacterial mRNA, cDNA library, trpA mRNA, lpp mRNA.

1. Introduction

Poly(A) sequences at the 3′-end have long been recognized as a characteristic structural feature of eucaryotic mRNA. Nevertheless, the function of poly(A) sequences has not yet been elucidated in molecular terms, although various possibilities have been suggested[1]. Until recently, it was generally accepted that post-transcriptional reactions leading to 3′-polyadenylation of mRNA did not occur in bacterial cells. However, through the development of an improved RNA isolation procedure, it was possible to demonstrate that large amounts of polyadenylated RNA exist also in *E. coli*, *B. subtilis*, and *B. brevis*[2]. Characterization of bacterial poly(A)RNA has revealed much shorter poly(A) tracts than in eucaryotes, associated with metabolically unstable molecules that are functional mRNAs[2–4]. More recently, it has been possible to identify poly(A) sequences at the 3′-termini of specific bacterial mRNA molecules, leaving no doubt that polyadenylation is an important attribute of procaryotic as well as of eucaryotic mRNA[5,6].

2. Detection of poly(A)RNA in different bacterial species

Making use of our one-step RNA isolation procedure, involving lysis by proteinase K in the presence of sodium dodecyl sulfate, and adsorption to oligo(dT)-cellulose, we have been able to detect substantial amounts of poly(A)RNA in pulse-labeled RNA of *B. subtilis* and *E. coli* as well as *B. brevis*[2]. As shown in Table I, about 20% of pulse-labeled RNA was polyadenylated, whereas stable RNA species were devoid of poly(A).

Our result contrasts with that of earlier workers[7] who found much lower levels of poly(A)RNA in *E. coli*. This discrepancy is due to the fact that bacterial poly(A)RNA is selectively lost during the phenol extraction procedure customarily used for RNA isolation but avoided in our studies[2].

Table I
Proportion of poly(A)-RNA in pulse-labeled RNA from different bacterial species

Organism	Stage of growth	Labeled precursor	% poly(A)-RNA[a]
B. brevis	Exponential	^{3}H-Ado	38
B. brevis	Exponential	^{3}H-Urd	30
B. brevis	Sporulating	^{3}H-Ado	37
B. brevis	Sporulating	^{3}H-Urd	26
B. subtilis	Exponential	^{3}H-Ado	16
B. subtilis	Exponential	^{3}H-Ado	0.2[b]
B. subtilis	Sporulating	^{3}H-Ado	12
E. coli	Exponential	^{3}H-Ado	24
E. coli	Exponential	^{3}H-Urd	18

[a] Per cent of radioactive RNA adsorbed to oligo(dT)-cellulose at high ionic strength and eluted by water.
[b] A 1-hour incubation with unlabeled adenosine (0.4 mM) followed the pulse-labeling period.

3. Characterization of poly(A)RNA from bacteria

Fifteen to twenty per cent of pulse-labeled RNA in *B. subtilis* was found to bind to oligo(dT)-cellulose at high ionic strength, an indication of the presence of poly(A) sequences. After treatment with pancreatic and T1 ribonucleases, 10–15% of purified [^{3}H]adenosine-labeled poly(A)RNA but only 0.5% of [^{3}H]uridine-labeled poly(A)RNA retained the ability to bind to oligo(dT)-cellulose. This indicated that the poly(A) sequences were covalently linked to other RNA sequences and constituted about 3–5% of the total nucleotide residues.[4]

In order to determine the position of the poly(A) tracts in the RNA molecules, [^{3}H]adenosine-labeled RNA was treated with pancreatic and T1 ribonucleases, and the resulting poly(A) tracts were digested with snake venom phosphodiesterase in the absence

and in the presence of bacterial alkaline phosphatase. In the absence of alkaline phosphatase, 50–65% of the poly(A) tracts were digested, indicating that they were originally located at the 3'-end of the RNA molecule. On the other hand, the remaining 35–50% of the poly(A) tracts must have been separated from the 3'-terminus by at least one nucleotide residue. These results are very similar to those obtained with poly(A)RNA from *B. brevis*[3].

The size distribution of the poly(A) tracts obtained by RNase T1 and A digestion of *B. subtilis* RNA was determined by polyacrylamide gel electrophoresis. The weight-average length was 49 nucleotide residues and the number-average length 37 residues, values that were about 20% lower than those found with *B. brevis*[3,4]. We compared the ability of different RNA fractions to serve as a template for protein synthesis in the cell-free system from *E. coli* MZ9. As shown in fig. 1, poly(A)RNA was about eight times more effective in stimulating [^{3}H]leucine incorporation than unfractionated RNA or non-poly(A)RNA[4]. Hybridization experiments using the method of Southern have shown that many different types of RNA sequences are polyadenylated, and also that most poly(A)RNA species have an unadenylated counterpart[4]. Our earlier studies discovered a c-GMP-sensitive intracellular 3-exonuclease in *B. brevis* as a major factor in mRNA degradation[8]. We found that non-poly(A)RNA was rapidly hydrolyzed by the 3'-exonuclease whereas poly(A)RNA was quite resistant, indicating some link of mRNA stability to polyadenylation[9].

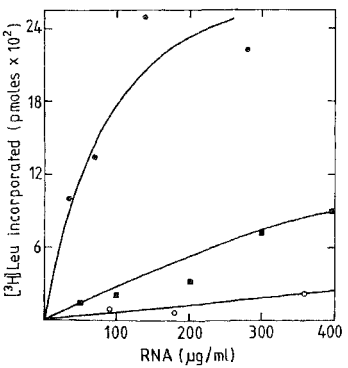

FIG. 1. Effect of RNA concentration on [^{3}H]leucine incorporation into protein in an *E. coli* cell-free system by using various types of RNA from *B. subtilis*. Total RNA, poly(A)RNA, and non-poly(A)RNA were isolated, messenger activity of these RNA fractions was assayed by using *E. coli* MZ9 extract, and [^{3}H]leucine incorporation into protein was measured as described. Poly(A)RNA (●), total RNA (■), and non-poly(A)RNA(○).

4. Construction of a cDNA library from poly(A)RNA of *B. subtilis*

We have shown in fig. 2 that poly(A)RNA from *B. subtilis* can serve as a template for the synthesis of complementary DNA by reverse transcriptase using oligo(dT) as primer[10]. The cDNA thus synthesized can serve as template for double-stranded cDNA synthesis. The ds cDNA could be inserted into the PstI site of pBR322 and cloned in *E. coli* DH1. The cDNA inserts from a few cloned recombinant pBR322 plasmids were transferred to M13mp18 bacteriophage for sequence determination[11]. DNA sequences of the sense strand for six such clones are shown in fig. 3. The length of the poly(A) sequence observed ranged from 4–19 adenylate residues and was followed by at least one translation termination codon in one of the possible reading frames. A search for common elements in the six cloned cDNA segments revealed no sequences that could give rise to a consistent pattern of RNA secondary structure. Indeed, a striking aspect of the DNA sequences was the paucity of elements that could give rise to RNA secondary structure, with no regions of dyad symmetry involving more than five contiguous base pairs. On the other hand, as indicated by underlining in fig. 3, five of the six sequences had in common the element $T^{C}_{G}GTC$, mostly far upstream from the start of the poly(A) tract, and four of the six had the sequence ATT^{CC}_{GG} just adjacent to the poly(A). It is of interest that the latter sequence (ATT^{CC}_{GG}) occurs frequently in the 3′-untranslated region of a large number of *E. coli* and *B. subtilis* operons and may perhaps have a role in RNA processing. In conclusion, our results constitute

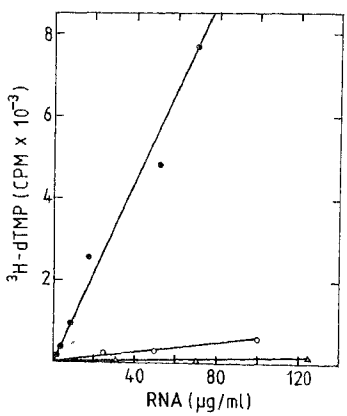

FIG. 2. Comparison of template activity for cDNA synthesis for various RNA fractions. Standard reactions were done to measure [^{3}H]dTMP incorporation into cDNA using varying amounts of different RNA. Poly(A)RNA (●), total RNA (○), and ribosomal RNA (△).

FIG. 3. Partial nucleotide sequence of the oligo(adenylate)-containing terminal region of *B. subtilis* cDNA cloned in M13mp18. The sequences shown are those of the strands with oligo(dA) adjacent to the oligo(dC)-termini and the numbers indicate the distance from the start of the oligo(dA) tract. The putative consensus sequences ATT^{GG}_{GC} and $T^{G}_{C}GTC$ are indicated by underlining or asterisks, respectively. The potential translation termination codons closest to the oligo(dA) tract that define an uninterrupted reading frame are indicated by italics.

independent and direct evidence for the existence of bacterial poly(A)-containing mRNA. This is the first report on the construction of cDNA library in procaryotes.

5. Polyadenylation of specific messages in *E. coli*

To elucidate the biological function of polyadenylation, we have undertaken a systematic study of polyadenylation of several specific mRNAs from *E. coli*. We studied in detail two mRNAs, those encoded by the *lpp*[5] and *trpA*[6] genes. *lpp* is a monocistronic gene, expressed constitutively, represents the most abundant and one of the most stable mRNAs in *E. coli* and its product is a structural protein that is secreted through the cytoplasmic membrane. On the other hand, *trpA* gene is a part of a polycistronic biosynthetic operon under repression and transcription attenuation control and has a short half life typical of most *E. coli* mRNAs. These studies involved purification of specific mRNA from *E. coli* transformed with a plasmid carrying the specific gene which is deleted from the chromosome. Several lines of evidence: (i) isolation of poly(A) tract after digestion with RNase A and T1, (ii) depolyadenylation by treatment with RNase H and oligo(dT), and (iii) synthesis of cDNA by reverse transcriptase with $oligo(dT)_{12}$ and specific mRNA as template indicated that both mRNAs are extensively polyadenylated. In view of the quite different functions

Table II
Comparison of polyadenylation of trpA mRNA with lpp mRNA

Type of RNA	Degree of polyadenylation of			
	trpA mRNA		lpp mRNA	
	% Poly(A)-RNA	Poly(A) length (bp)	% Poly(A)-RNA	Poly(A) length (bp)
Steady state	36	9	47	9
Newly synthesized	48	15–20	42	10–15
In vitro synthesized	45	14	50	15

of these mRNAs, it is of interest to compare their levels of polyadenylation (Table II). The values obtained for the poly(A) length and poly(A)RNA content of the two different messages were very similar in both newly synthesized or steady-state levels. The striking similarities indicate that polyadenylation of mRNA may be a general phenomenon in *E. coli*, unrelated to the specific physiological role of the mRNA. Using the same experimental approaches, we have been able to show that many other genes, *e.g.* alkaline phosphatase, *tet* gene, ampicillin and ompA gene, have polyadenylated mRNA. This is also true for several genes, *e.g.* glutaimine synthetase, aspartokinase II, DNA gyrase, recF, metC, in *B. subtilis.*

6. Biosynthesis of poly(A) at the 3′-end of mRNA

In view of the fact that the *trpA* and *lpp* genes and their 3′-untranslated region contain no continuous deoxyadenylate sequences larger than five nucleotides, one can conclude that the polyadenylate moiety is added post-transcriptionally. To study the enzymatic process of polyadenylation, eariler we used permeable *E. coli* cells which are capable of synthesizing poly(A)RNA[12]. Using permeable cells of *E. coli* with a deletion in the chromosomal *lpp* or *trpA* gene but transformed with a plasmid carrying the desired specific gene, we were also able to synthesize polyadenylated *lpp* mRNA or *trpA* mRNA[5,6]; the maximum length of poly(A) tracts associated with the specific mRNA *in vitro* was 14 nucleotide residues.

7. Conclusion

We have made considerable progress in demonstrating unambiguously the presence of polyadenylated RNA in *B. subtilis* and *E. coli* and showing that poly(A) sequences are associated with defined mRNA species. In the light of this fact, it is apparent that the process of transcription termination is closely interlined with polyadenylation and various kinds of 3′-terminal processing of mRNA in procaryotes. Extensive studies in all these areas may clarify the picture in the future.

Acknowledgement

This work was supported by grants from the National Institutes of Health, GM26517, and the Biomedical Research Support Program, PR05711, of the Public Health Service.

References

1. LITTAUER, U. Z. AND SOREQ, H. — The regulatory function of poly(A) and adjacent 3' sequences in translated RNA, *Prog. Nucl. Acid Res. Mol. Biol.*, 1982, Vol. 27, pp. 53–83, Academic Press.

2. GOPALAKRISHNA, Y., LANGLEY, D. AND SARKAR, N. — Detection of high levels of polyadenylate-containing RNA in bacteria by the use of a single-step RNA isolation procedure, *Nucl. Acids Res.*, 1981, **9**, 3545–3554.

3. SARKAR, N., LANGLEY, D. AND PAULUS, H. — Isolation and characterization of polyadenylate-containing RNA from *Bacillus brevis*, *Biochemistry*, 1978, **17**, 3468–3474.

4. GOPALAKRISHNA, Y. AND SARKAR, N. — Characterization of polyadenylate-containing ribonucleic acid from *Bacillus subtilis*, *Biochemistry*, 1982, **21**, 2724–2729.

5. TALJANIDISZ, J., KARNIK, P. AND SARKAR, N. — Messenger ribonucleic acid for the lipoprotein of the *Escherichia coli* outer membrane is polyadenylated, *J. Mol. Biol.*, 1987, **193**, 507–515.

6. KARNIK, P., TALJANIDISZ, J., SASVARI-SZEKELY, M. AND SARKAR, N. — 3-Terminal polyadenylate sequences of *Escherichia coli* tryptophan synthetase α-subunit messenger RNA, *J. Mol. Bio.*, 1987, **196**, 347–354.

7. NAKAZATO, H., VENKATESAN, S. AND EDMONDS, M. — Polyadenylic acid sequences in *E. coli* messenger RNA, *Nature*, 1975, **256**, 144–146.

8. SARKAR, N. AND PAULUS, H. — A guanosine 3':5'-monophosphate-sensitive nuclease from *Bacillus brevis*, *J. Biol. Chem.*, 1975, **250**, 684–690.

9. GOPALAKRISHNA, Y. AND SARKAR, N. — Selective resistance of bacterial polyadenylate-containing RNA to hydrolysis by guanosine 3'–5'-monophosphate-sensitive nuclease of *Bacillus brevis*, *Biochem. Biophys. Res. Commun.*, 1981, **103**, 454–460.

10. GOPALAKRISHNA, Y. AND SARKAR, N. — The synthesis of DNA complementary to polyadenylate-containing RNA from *Bacillus subtilis*, *J. Biol. Chem.*, 1982, **257**, 2747–2750.

11. KARNIK, P., GOPALAKRISHNA, Y. AND SARKAR, N. — Construction of a cDNA library from polyadenylated RNA of *Bacillus subtilis* and the determination of some 3'-terminal sequences, *Gene*, 1986, **49**, 161–165.

12. GOPALAKRISHNA, Y. AND SARKAR, N. — Synthesis of polyadenylate-containing RNA *in vitro* in permeable cells of *Escherichia coli* B, *Arch. Biochem. Biophys.*, 1983, **224**, 196–205.