Annu. Rev. Biochem. 1997. 66:173–97
Copyright © 1997 by Annual Reviews Inc. All rights reserved

POLYADENYLATION OF mRNA IN PROKARYOTES

Nilima Sarkar
Boston Biomedical Research Institute, Boston, Massachusetts 02114 and Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115; e-mail: sarkar@bbri.harvard.edu

KEY WORDS: poly(A) RNA, poly(A) polymerase, bacterial cDNA libraries, mRNA processing, mRNA degradation

ABSTRACT

The 3′-ends of both prokaryotic and eukaryotic mRNA are polyadenylated, but the poly(A) tracts of prokaryotic mRNA are generally shorter, ranging from 15 to 60 adenylate residues and associated with only 2–60% of the molecules of a given mRNA species. The sites of polyadenylation of bacterial mRNA are diverse and include the 3′-ends of primary transcripts, the sites of endonucleolytic processing in the 3′ untranslated and intercistronic regions, and sites within the coding regions of mRNA degradation products. The diversity of polyadenylation sites suggests that mRNA polyadenylation in prokaryotes is a relatively indiscriminate process that can occur at all mRNA's 3′-ends and does not require specific consensus sequences as in eukaryotes. Two poly(A) polymerases have been identified in *Escherichia coli*. They are encoded by unlinked genes, neither of which is essential for growth, suggesting significant functional overlap. Polyadenylation promotes the degradation of a regulatory RNA that inhibits the replication of bacterial plasmids and may play a similar role in the degradation of mRNA. However, under certain conditions, poly(A) tracts may lead to mRNA stabilization. Their ability to bind S1 ribosomal protein suggests that poly(A) tracts may also play a role in mRNA translation.

CONTENTS

INTRODUCTION 174
DETECTION OF POLY(A) RNA IN BACTERIA 175
Association of Poly(A) Tracts with mRNA 175
Poly(A) Tracts Associated with Specific mRNAs 175
STRUCTURE OF BACTERIAL POLY(A) RNA 176
Bacterial cDNA Libraries 176

0066-4154/97/0701-0173$08.00

PARSE EX.1082.001

Sites of Polyadenylation of Specific mRNAs 177
SYNTHESIS OF BACTERIAL POLY(A) RNA 182
Bacterial Poly(A) Polymerases 182
Poly(A) Polymerase I 183
Poly(A) Polymerase II 183
Other Bacterial Poly(A) Polymerases 185
In Vitro Poly(A) RNA Synthesis 185
POSSIBLE FUNCTIONS OF RNA POLYADENYLATION IN BACTERIA 186
Modulation of the Control of Plasmid Regulation by Antisense RNA 186
Relationships Between Poly(A) Polymerases and 3′-Exonucleases and Their Effect on mRNA Half-Life 187
Involvement in Translation of mRNA 190
mRNA POLYADENYLATION IN EUKARYOTIC ORGANELLES 190
Mitochondria 191
Chloroplasts 192
mRNA POLYADENYLATION IN ARCHAEA 193
CONCLUDING REMARKS 193

INTRODUCTION

Many of the seminal studies on the mechanisms of gene expression and macromolecular biosynthesis in the 1960s employed bacterial model systems with the tacit assumption that the processes within these systems are universal. The research group at the Institute Pasteur was especially sanguine in this respect, as evidenced by Jacques Monod's dictum: "What is true for *E. coli* is true for an elephant." Yet careful reflection suggests that there should be significant mechanistic differences between prokaryotes and eukaryotes, owing not least to the subdivision of the eukaryotic cell by intracellular membrane systems. These differences are most clear in mRNA metabolism. In prokaryotes, mRNA is accessible to the protein synthetic machinery as soon as it is transcribed from DNA. Because translation is essentially cotranscriptional, bacterial mRNA is utilized very rapidly and efficiently, especially because many genes can be transcribed together in a single polycistronic mRNA, and generally has a relatively short half-life. On the other hand, although most mRNAs are translated in the cytoplasm, transcription in eukaryotes occurs in the nucleus, a circumstance that imposes spatial and temporal barriers between the transcription and translation processes; these barriers are further enhanced by the necessity of splicing intervening sequences from most eukaryotic transcripts before they can be translated. As a result, the primary transcripts, which generally encode only single proteins, must be processed and translocated across the nuclear membrane prior to translation and therefore have relatively long half-lives. Except for RNA splicing, the function of eukaryotic RNA processing, which involves modification of the 5′-ends by capping and of the 3′-ends by polyadenylation, is not entirely clear and may be related to transport from the nucleus to the cytoplasm, to translation by the ribosome, or to mRNA turnover. Because not all of these processes have

PARSE EX.1082.002

counterparts in prokaryotic mRNA metabolism, it was long thought that mRNA processing events such as polyadenylation were unique to eukaryotes. Therefore, when poly(A) tracts were first found in bacterial mRNA, they were considered to be biochemical curiosities rather than general attributes of bacterial mRNA. This skepticism was not entirely unjustified in view of the difficulties associated with the characterization of the rather unstable bacterial mRNA.

This review summarizes current knowledge of mRNA polyadenylation in prokaryotes. An earlier review (1) provides a summary of the biochemical studies that led to the recognition of bacterial poly(A) RNA. The present review concentrates on the more recent molecular and genetic studies that address the structure, synthesis, and function of poly(A) RNA in prokaryotes, with emphasis on work published since 1985 that deals primarily with *E. coli.* However, to provide an evolutionary perspective on mRNA polyadenylation in prokaryotes and eukaryotes, I also summarize what is known about this process in eukaryotic organelles, which originated as prokaryotic endosymbionts, and in the third kingdom of life, the archaea.

DETECTION OF POLY(A) RNA IN BACTERIA

Association of Poly(A) Tracts with mRNA

Although a poly(A) polymerase was purified from *E. coli* as early as 1962 (2), it was not until 1975 that bacterial poly(A) RNA could be clearly identified. Studies that took advantage of the specific binding of poly(A) tracts to oligo(dT) cellulose showed that they are associated with pulse-labeled RNA of *E. coli* (3, 4) and *Caulobacter crescenta* (5). In the next few years, polyadenylated RNAs were also identified in *Bacillus brevis* (6), *Hyphomicrobium* (7), *Bacillus subtilis* (8, 9), and in two filamentous cyanobacteria (10). The association of poly(A) tracts with functional mRNAs was demonstrated by the use of in vitro translation assays in poly(A) RNA isolated from *B. brevis* (11), *B. subtilis* (9), *Rhodopseudomonas capsulata* (12), and *Rhodospirillum rubrum* (13). The general consensus that emerged from these studies with total bacterial RNA is that (*a*) the poly(A) tracts of bacterial mRNAs range between 14 and 60 nucleotides in length, which is significantly shorter than the 80–200 adenylates associated with eukaryotic mRNAs; and (*b*) only a relatively small fraction of mRNA molecules (1–40%, depending on the source and method of RNA analysis) are polyadenylated, in contrast to the virtually quantitative polyadenylation of most eukaryotic mRNAs.

Poly(A) Tracts Associated with Specific mRNAs

The length of poly(A) tracts and the degree of polyadenylation of total mRNA are average values, which are difficult to interpret. It was therefore important

PARSE EX.1082.003

to characterize the poly(A) tracts associated with specific mRNAs. Two genetically and biochemically well characterized *E. coli* mRNAs, which cover a broad range of structure and function, were chosen as subjects for investigation (14, 15). The first was the mRNA for the outer membrane lipoprotein, which is encoded by *lpp*, the sole gene in a constitutive monocistronic operon. *lpp* RNA is the most abundant and one of the most stable mRNAs in *E. coli*, whose product is a structural protein that is secreted through the cytoplasmic membrane. The second was the mRNA for the α-subunit of tryptophan synthase, which is encoded by *trpA*, the distal gene in a polycistronic biosynthetic operon under repression and transcription attenuation control, and has a short half-life typical of most *E. coli* mRNAs. Both mRNAs were expressed in *E. coli* from multicopy plasmids, and isolated and identified by hybrid selection with appropriate single-stranded DNA probes. The association of poly(A) tracts with *lpp* and *trpA* mRNAs was demonstrated by their ability to bind to oligo(dT) cellulose, unless they are first treated with oligo(dT) and RNase H to remove poly(A) tracts, and by their ability to serve as templates for the oligo(dT)-dependent synthesis of cDNA by reverse transcriptase. Parallel studies were carried out on newly synthesized mRNA identified by in vivo pulse labeling; on the steady state mRNA population, isolated by hybrid selection with specific gene probes followed by 3′-end labeling; and on RNA synthesized in permeable cells, described in the section on In Vitro Poly(A) RNA Synthesis. The results obtained with *lpp* mRNA (14) and with *trpA* mRNA (15) are strikingly similar: about 40–50% of the mRNAs were polyadenylated, with poly(A) tracts of about 10–20 adenylate residues in both the newly synthesized and the steady state RNA fractions. The striking similarities between these diverse mRNAs suggest that polyadenylation of *E. coli* mRNA is a general phenomenon, unrelated to the specific physiological role of the mRNA.

STRUCTURE OF BACTERIAL POLY(A) RNA

Bacterial cDNA Libraries

An important advance in the study of bacterial poly(A) RNA, which paved the way for subsequent analysis at the molecular level, was the oligo(dT)-dependent synthesis of DNA complementary to bacterial mRNA (cDNA) by reverse transcriptase. Using purified poly(A) RNA from *B. subtilis* as a template, the average chain length of the cDNA was found to be 400 nucleotides, with a range of 230–800 (16). The cDNA thus synthesized could serve as a template for double-stranded cDNA synthesis, making possible its insertion into plasmids such as pBR322 and subsequent cloning in *E. coli* to yield a cDNA library (17). Six clones from the first such cDNA library (400 clones) were sequenced and found to carry runs of 4–19 deoxyadenylate residues that are 1–90 nucleotides

PARSE EX.1082.004

downstream of the translation termination codon of a potential open reading frame. The lengths of the poly(dA) tracts in the cDNA do not necessarily reflect the poly(A) lengths in the corresponding mRNA but rather that of the oligo(dT) primers used in cDNA synthesis, and they may have been further shortened as a result of the extensive S1 nuclease treatment that was needed to allow the cloning of double-stranded cDNA. Although no specific genes were identified in the *B. subtilis* cDNA library at that time, this approach was subsequently applied to mRNA from *E. coli* and has led to the molecular characterization of the polyadenylation sites in several mRNAs. A cDNA library has recently been prepared from *E. coli* but has not yet been analyzed in detail (18).

Sites of Polyadenylation of Specific mRNAs

Specific cDNAs have been cloned with the use of oligo(dT) primers with a 5′-adapter that could serve as a primer for the polymerase chain reaction (PCR). In conjunction with a second PCR primer corresponding to a portion of the coding region of the mRNA of interest, the cDNA was amplified and then cloned into appropriate plasmids for subsequent nucleotide sequence analysis. Owing to the rapid turnover of poly(A) tracts in wild-type *E. coli*, the successful application of this procedure to *E. coli* required using RNA from strains with defects in the *pnp* and *rnb* loci, encoding the major 3′-exonucleases, polynucleotide phosphorylase (PNPase), and RNase II, respectively. A few studies employed *E. coli* strains with an additional lesion in *rne*, the gene for the endonuclease RNase E. As noted below, the use of such mutants may affect the relative amounts of the various types of poly(A) RNA molecules detected.

The application of this type of experimental approach to various bacterial mRNAs has uncovered six different classes of mRNA polyadenylation sites, which are summarized schematically in Figure 1. The individual poly(A) RNAs studied are described in more detail below.

lpp mRNA The *lpp* transcript is one of the most abundant and stable mRNAs in *E. coli* and the only bacterial mRNA that has been purified and sequenced (19). It is a monocistronic mRNA that is terminated by the stem-loop structure characteristic of rho-independent transcription terminators. Analysis of *lpp* mRNA from a *pnp rnb* strain of *E. coli* yielded four clones corresponding to polyadenylated *lpp* transcripts (20). In one of these, the poly(A) tract was attached to the end of the primary transcript defined by the rho-independent transcription terminator, corresponding to Class I poly(A) RNA listed in Figure 1. In three other clones, the mRNA was truncated and the poly(A) moiety attached just upstream of the terminal stem-loop structure at a site that is a candidate for endonucleolytic cleavage by enzymes such as RNase E [Figure 1; Class II poly(A) RNA]. The possibility that this second class of poly(A) RNA is an artifact caused by

PARSE EX.1082.005

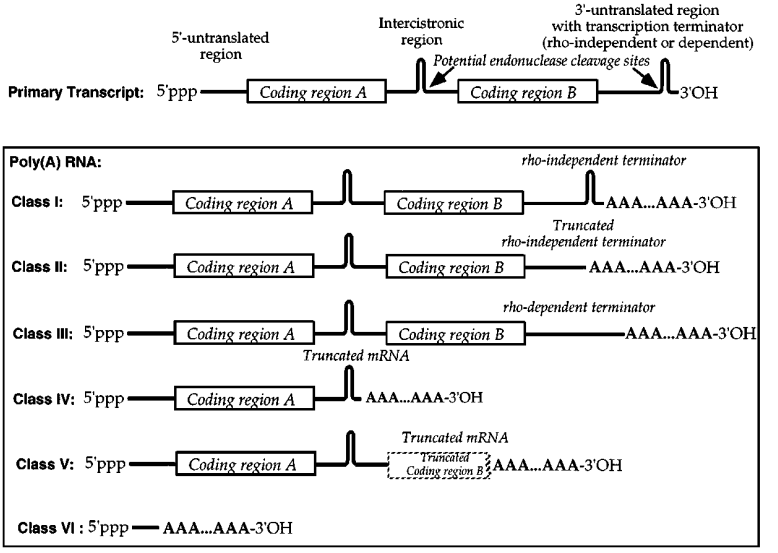

Figure 1 Classification of the sites of mRNA polyadenylation observed in bacteria.

oligo(dT) priming at a purine-rich internal region was eliminated by the observation that some of the poly(A) tracts found were longer than the oligo(dT) primer. An important implication of the existence of Class II poly(A) RNA is that mRNA polyadenylation is not necessarily coupled to transcription but can occur after endonucleolytic processing of the primary transcript. On the other hand, the observation that *lpp* clones corresponding to Class II poly(A) RNA were three times more frequent than Class I clones is not necessarily an indication of the relative amounts of these two classes of poly(A) RNA in wild-type *E. coli* but may reflect the persistence of mRNA molecules partially degraded by endonucleases in the absence of PNPase and RNase II.

rpsO mRNA The *rpsO* mRNA encoding ribosomal protein S15 was studied in strains of *E. coli* with defects in various ribonucleases (21, 22). Disregarding clones derived from read-through mRNA that failed to terminate at the rho-independent termination site and artifacts caused by oligo(dT)-priming at an A-rich sequence in the *rpsO* coding region, the following distributions of cloned cDNAs were obtained. In *pnp rnb* strains lacking the two major 3′-exonucleases, only 3% Class I poly(A) RNA was obtained compared to 76% Class II poly(A) RNA, a trend similar to that shown by *lpp* mRNA in the same *E. coli* strain (20). In contrast, with *pnp rnb rne* triple mutants, which lack the

PARSE EX.1082.006

endonuclease RNase E in addition to the major 3′-exonucleases, 64% Class I and only 7% Class II poly(A) RNA was found, suggesting that RNase E plays a key role in the posttranscriptional cleavage of the terminal stem-loop structure, which is a prerequisite for polyadenylation at Class II sites. Besides Class I and Class II poly(A) RNA, both types of 3′-exonuclease-deficient strains had a third class of cDNA, accounting for about 25% of the cDNA clones, in which the poly(A) tracts were attached to a truncated *rpsO* coding region (21). This type of poly(A) RNA, which is referred to as Class V poly(A) RNA in Figure 1, may not be a major constituent in wild-type *E. coli*, where mRNA molecules that suffer endonucleolytic cleavage within coding regions are probably rapidly degraded by 3′-exonucleases. Such an interpretation is supported by the observation that strains with normal 3′-exonucleases (pnp^+ rnb^+ *rne*) have 96% Class I and only 4% Class V poly(A) RNA (21). Notwithstanding its probable lack of physiological significance, the presence of Class V poly(A) RNA in the 3′-exonuclease-deficient *E. coli* strains suggests that mRNA polyadenylation is a relatively indiscriminate process that occurs at any 3′-end of an mRNA molecule, irrespective of its origin.

crp mRNA The nucleotide sequence just downstream of the *E. coli crp* locus (23, 24), which encodes the cyclic AMP receptor protein, includes a potential stem-loop structure resembling rho-independent transcription terminators, and further downstream, a sequence enriched in C residues (34%) and deficient in G (18%), characteristic of rho-dependent transcription terminators (25). This arrangement of potential transcription terminator sites is similar to the disposition of the rho-independent (*trp* t) and rho-dependent (*trp* t′) terminator sites at the end of the *E. coli trp* operon (26). In a cDNA clone derived from a *pnp rnb rne* strain of *E. coli*, the site of polyadenylation was found to be 106 nucleotides downstream of the translation termination site, which is at the distal end of the putative rho-dependent transcription terminator (27, 28). This finding suggests that mRNA polyadenylation can occur also at rho-dependent transcription termination sites, leading to an additional category of poly(A) RNA, referred to as Class III poly(A) RNA in Figure 1.

lacZ mRNA Although the *E. coli lacZ* gene encoding β-galactosidase is the first cistron in the *lacZYA* operon, a significant fraction of *lac* mRNA terminates in the *lacZY* intercistronic region, which contains a stem-loop structure that may be the site of posttranscriptional endonucleolytic cleavage (29). Most *lacZ*-specific cDNA clones were found to have poly(A) tracts just distal to the intercistronic stem-loop structure, regardless of whether cDNA was derived from wild-type *E. coli* or from strains lacking PNPase, RNase II, RNase E, or combinations of these enzymes (30). This pattern of polyadenylation, with poly(A) tracts at

PARSE EX.1082.007

sites of endonucleolytic processing in intercistronic regions, leads to a functionally distinct class of poly(A) RNAs referred to as Class IV poly(A) RNA in Figure 1.

rho mRNA The synthesis of the transcription termination factor rho, encoded by the *E. coli rho* gene, is autogenously controlled by transcription attenuation at two rho-dependent premature transcription termination sites (A1 and A2) in the untranslated leader region (31). The *trxA* and *rho* genes are adjacent on the *E. coli* chromosome, with the rho-independent transcription termination site of *trxA* being just upstream of the transcription initiation site of *rho*, and there is evidence for some transcriptional read-through to yield a bi-cistronic mRNA (31). Several *trxA*-specific cDNA clones derived from a *pnp rnb rne* strain of *E. coli* have poly(A) tracts at either of the rho-dependent *rho* transcription attenuator sites (G-J Cao & N Sarkar, unpublished observations). One of the polyadenylation sites, at residue 162, corresponds to a transcription termination site found in an in vitro transcription assay using *rho* as template in the presence of Rho protein (DF Bear, submitted for publication), suggesting that rho-dependent transcription attenuation sites are subject to polyadenylation. Assuming that polyadenylation at the transcription attenuation sites occurs not only in *trx-rho* read-through transcripts but also in transcripts starting at the *rho* promoter, this would define a new class of poly(A) RNAs, Class VI poly(A) RNA (Figure 1). An interesting question, which has not yet been addressed, is whether leader transcripts terminating at the much more common rho-independent transcription attenuator sites of *E. coli* biosynthetic operons (32) are also polyadenylated.

OTHER *E. coli* mRNAs Several cDNA clones corresponding to *lacA* mRNA from wild-type *E. coli* as well as from *pnp*, *rnb*, and *rne* mutants have been sequenced (30). *lacA* is the distal gene in the lactose operon, and its coding sequence is followed by two potential stem-loop structures separated by about 90 nucleotides, but it has not been established whether either or both of these function as rho-independent transcription terminators (29). All *lacA* cDNAs analyzed were polyadenylated at sites just distal to these stem-loop structures and therefore correspond to Class I or Class II poly(A) RNAs, depending on whether the stem loops are transcription termination or posttranscriptional processing sites, respectively.

The transcription of the *trxA* gene, encoding thioredoxin, is terminated by a rho-independent mechanism involving a stem-loop structure just downstream of the coding region. A *trxA* cDNA clone, derived from mRNA isolated from a *pnp rnb rne E. coli* strain using a specific *trxA* PCR primer, was found to be polyadenylated just upstream of the rho-independent termination site, consistent with Class II poly(A) RNA, suggesting that some cleavage of the terminal

PARSE EX.1082.008

stem-loop structure can occur in the absence of RNase E (G-J Cao & N Sarkar, unpublished observations). On the other hand, an *E. coli* cDNA library prepared from the same *E. coli* strain yielded two clones that were identified with known mRNA sequences, *cysZ* and *ndaA*, both of which were polyadenylated just distal to the terminal stem loops and thus represented Group I poly(A) RNA (G-J Cao & N Sarkar, unpublished observations).

B. subtilis hag mRNA Although the first bacterial cDNA library was prepared from *B. subtilis* mRNA (17), most subsequent studies were focused on *E. coli*. cDNA clones have been isolated for only one specific *B. subtilis* mRNA, that corresponding to the *hag* gene, which encodes flagellin (33). Owing to the intrinsically higher levels of poly(A) RNA in *B. subtilis* than in *E. coli*, the mRNA for these experiments could be isolated from wild-type bacteria rather than from ribonuclease-deficient mutant strains. In all six cDNA clones that were sequenced, the site of polyadenylation was just upstream of the stem-loop structure of the rho-independent termination, indicating that polyadenylation occurred subsequent to posttranscriptional processing of the primary transcript (33). The occurrence of Group II poly(A) RNA in wild-type *B. subtilis* suggests that this type of polyadenylation is physiologically significant and not an artifact caused by the absence of 3′-exonucleases.

BACTERIOPHAGE T7 mRNAs The transcription of bacteriophage T7 DNA involves two different RNA polymerases; the "early" genes are transcribed by *E. coli* RNA polymerase, and the "middle" and "late" genes are transcribed by T7 RNA polymerase, both processes yielding polycistronic transcripts, which are processed at specific sites by RNase III (see e.g. 34). Studies by Kramer et al (35) revealed the presence of one or more uncoded adenylate residues at the 3′-ends of early mRNA, and it was therefore of interest to determine whether these residues derived from poly(A) tracts and whether all or only some of the bacteriophage T7 mRNA molecules were polyadenylated. Screening of pulse-labeled RNA fractions corresponding to early, middle, and late transcripts for poly(A) tracts by oligo(dT)-cellulose chromatography showed about 30% of the RNA in each fraction to be polyadenylated (MD Johnson & N Sarkar, unpublished data). Analysis of selected clones from a bacteriophage T7 cDNA library, prepared from pooled early, middle, and late RNA and screened with specific bacteriophage T7 gene probes, showed extensive uncoded 3′-terminal extensions that were highly enriched in adenylate residues (60–100%) (36). In four clones (genes 0.3, 0.4, 0.5, and 6.5), the sites of polyadenylation were in the intercistronic regions and therefore representative of Class IV poly(A) RNA (Figure 1) (36). Of special interest was the gene 0.4 clone, where the polyadenylation site corresponded exactly to the site of RNase III cleavage. In

PARSE EX.1082.009

three other clones, the polyadenylation sites were within the coding regions, as in Class V poly(A) RNA (Figure 1) (36).

Bacteriophage T7 mRNA can thus be polyadenylated regardless of whether it is transcribed by the host- or bacteriophage-specific RNA polymerase, with the polyadenylation sites either at the site of RNase III processing or at secondary cleavage sites, including sites within coding regions. However, the poly(A) tracts associated with bacteriophage mRNA differ from those found in bacterial RNAs in that they are interspersed with up to 40% other nucleotides (36). As there is no evidence for a bacteriophage T7-specific poly(A) polymerase, the reduced nucleotide specificity of polyadenylation may be the result either of modification of the host poly(A) polymerases or of alterations in the sizes of the ribonucleoside triphosphate pools relative to ATP after bacteriophage infection. Poly(A) tracts interspersed with other nucleotides have also been found in chloroplast *pbsA* mRNA (37).

The polyadenylation of other bacteriophage mRNAs has not been investigated, but it should be noted that uncoded oligo(A) tracts have been found at the 3′-ends of bacteriophage λ transcripts, irrespective of rho-factor status (38), as well as in a small RNA (oop RNA) derived from the bacteriophage λ *ori* region (39).

COMPARISON WITH EUKARYOTIC POLY(A) RNA The great diversity in the polyadenylation sites found in bacterial mRNA, summarized in Figure 1, suggests that polyadenylation in bacteria is highly indiscriminate and occurs at virtually all unprotected RNA 3′-termini, irrespective of sequence and secondary structure. The promiscuous nature of the bacterial RNA polyadenylation process is strikingly illustrated by the Class V poly(A) RNAs—poly(A) tracts are found randomly distributed at the 3′-ends of what are evidently mRNA degradation products—and by the finding that some poly(A) tracts are associated even with ribosomal RNA (18). In contrast, mRNA polyadenylation in higher eukaryotes always occurs in the 3′ untranslated region of the processed transcript [analogous to Class II poly(A) RNA, shown in Figure 1], 10–30 nucleotides downstream of the hexanucleotide sequence AAUAAA (see e.g. 40). An exception is mitochondrial mRNA polyadenylation, discussed in the section on mRNA Polyadenylation in Eukaryotic Organelles, which also seems to be a relatively indiscriminate process.

SYNTHESIS OF BACTERIAL POLY(A) RNA

Bacterial Poly(A) Polymerases

Poly(A) polymerase (ATP:polyribonucleotide adenylyltransferase, EC 2.7.7.19) catalyzes the template-independent sequential addition of adenylate residues to

PARSE EX.1082.010

the 3′-hydroxyl termini of RNA molecules by the following reaction: RNA + nATP → $RNA(A)_n + nPP_i$. The first eukaryotic poly(A) polymerase was described in 1960 (41) and a corresponding bacterial enzyme in 1962 (2), considerably before the discovery of mRNA polyadenylation, but neither received much attention at that time. The only studies on bacterial poly(A) polymerases in the ensuing 15 years focused on the development of alternate purification procedures, which yielded enzyme preparations with different properties (42–44). It was only after considerable progress had been made in the characterization of bacterial poly(A) RNA that the focus of attention returned to bacterial poly(A) polymerases.

Poly(A) Polymerase I

The major poly(A) polymerase of *E. coli* (PAP I) was purified, and its gene was identified by amplifying and sequencing the coding region corresponding to its 25 amino-terminal amino acid residues (45). The 74-bp DNA sequence that was obtained matched a region of the *pcnB* locus of *E. coli*, a gene whose function was originally identified as controlling plasmid copy number (46); the gene was subsequently cloned and sequenced (47). It is interesting that the deduced translation start site is 17 amino acid residues upstream of the amino-terminal lysine residue of PAP I, which suggests that the enzyme is synthesized with a leader sequence that is subsequently lost (45). The coding region of the *pcnB* gene starts with the rarely used initiation codon UUG, which is also used by the *pnp* gene encoding polynucleotide phosphorylase (48), an enzyme that is also involved in the 3′ processing of mRNA. Direct evidence that the *pcnB* locus encodes poly(A) polymerase was provided by the observation that *E. coli* that is transformed with an inducible expression vector carrying *pcnB* produces 300-fold elevated levels of poly(A) polymerase upon induction (45). The deduced processed product of the *pcnB* gene is a 52-kDa protein with no significant homology with eukaryotic poly(A) polymerases but similar to a segment of *E. coli* tRNA nucleotidyltransferase that may be an RNA binding site (49). Even moderate overexpression of the *pcnB* gene is lethal to *E. coli* (45); on the other hand, disruption or deletion of *pcnB* reduced growth rates by less than 50% (47, 50). However, the replication of ColE1 plasmids, which requires the polyadenylation of RNAI, a regulatory antisense RNA, is suppressed in *pcnB* mutants (51). The polyadenylation of *rpsO* mRNA is also completely suppressed in *pcnB* deletion strains (22), but the average level of mRNA polyadenylation is reduced by less than 50% (52).

Poly(A) Polymerase II

The observation that substantial mRNA polyadenylation occurs even in *pcnB* deletion strains (52) suggests that *E. coli* has more than one poly(A) polymerase.

PARSE EX.1082.011

Indeed, a second poly(A) polymerase activity (PAP II) could be identified in extracts of *E. coli* with a *pcnB* deletion as a relatively hydrophobic protein with a molecular weight of about 35,000 (52).

Transposon disruption of a candidate gene for a second poly(A) polymerase (PAP II) was achieved by screening for *E. coli* mutants whose growth depends on the conditional expression of *pcnB*, a strategy that was predicated on the assumption that either PAP I or PAP II is required for viability (53). The coding region identified by this strategy was the open reading frame *f310*, located at 87 min on the *E. coli* chromosome. The deduced protein encoded by *f310* is relatively hydrophobic and has a molecular weight of 36,300, consistent with the properties of partially purified PAP II (52). Direct evidence that *f310* encodes PAP II was provided by overexpression of the gene, which led to the formation of inclusion bodies whose solubilization and renaturation yielded a 35-kDa protein with poly(A) polymerase activity; moreover, when the N-terminus of the *f310* coding region was fused to hexahistidine as an affinity-tag, a 36-kDa poly(A) polymerase activity could be purified on a Ni^{++} affinity column (53). *E. coli* PAP II has no significant sequence homology either to PAP I or to the viral and eukaryotic poly(A) polymerases, suggesting that the bacterial poly(A) polymerases have evolved independently.

An important question posed by the identification of the gene for a second *E. coli* poly(A) polymerase is whether there is functional overlap between PAP I and PAP II. It is clear that some RNA species, such as RNA I (51) and *rpsO* mRNA (22), depend solely on PAP I for polyadenylation. It is possible that PAP I and PAP II polyadenylate complementary classes of mRNA or that one of the enzymes initiates poly(A) chains and the other functions in their extension. However, it is more likely that there is significant functional overlap between PAP I and PAP II, because the deletion of the gene for either enzyme alone is not lethal (47, 50; G-J Cao & N Sarkar, unpublished observations). Indeed, redundancy with respect to enzymes acting on RNA seems to be almost the rule in *E. coli.* For example, 3′-exonucleolytic degradation of mRNA can be effected by both polynucleotide phosphorylase and RNase II, so mutational inactivation of either enzyme can be tolerated, but not the loss of both (54). Similarly, there is a high degree of functional overlap among the 3′-exoribonucleases involved in tRNA maturation, which include RNase D, RNase BN, RNase T, and RNase PH (55). The recently discovered *rnhB*, a gene encoding a second ribonuclease H (56), is another example of such functional duplication. Such redundancy may serve in part to guard against the loss of a vital function. Ironically, this functional and genetic redundancy has greatly hampered the genetic analysis of mRNA metabolism in *E. coli* and is primarily responsible for our ignorance of this vital metabolic area.

PARSE EX.1082.012

Although multiple poly(A) polymerases have also been described in eukaryotes (57), these are closely related proteins that are probably derived from a single gene by alternate splicing or posttranslational modification (58). In yeast, there appears to be only a single gene for poly(A) polymerase whose inactivation leads to the loss of viability (59).

Other Bacterial Poly(A) Polymerases

A gene whose product has significant homology to *E. coli* PAP I has been found in *B. subtilis* (60), but whether this gene encodes a functional poly(A) polymerase has not yet been tested. Other bacteria, such as *Mycobacterium leprae* (61) and *Haemophilus influenzae* (62), also contain coding regions for *E. coli* PAP I homologs. In addition, two unlinked open reading frames (sll0825 and sll1253) in the genome of the cyanobacterium *Synechocystis* sp. strain PCC6803 encode a 46-kDa and a 106-kDa protein, respectively, both of which are 40% identical with the deduced poly(A) polymerase sequence of *M. leprae* (63). On the other hand, the endoparasitic bacterium *Mycoplasma genitalium* has no open reading frames that encode proteins homologous to either PAP I or PAP II (64).

In Vitro Poly(A) RNA Synthesis

All in vitro studies on bacterial poly(A) polymerases have involved the use of unphysiological RNA primers, such as tRNA or poly(A), and can therefore provide little information on their in vivo function. For example, purified *E. coli* poly(A) polymerase can add more than 300 adenylate residues to tRNA (43), making it far longer than any bacterial poly(A) tract. Although differences in the preference of *E. coli* PAP I and PAP II for poly(A) and tRNA primers were found (52), it is difficult to interpret this observation because neither primer was a physiological one. In order to avoid these difficulties, poly(A) RNA synthesis has been studied in cells made selectively permeable to small molecules by treatment with toluene (65). Such an experimental system has the advantage that substrates and effectors can be manipulated with minimal disruption of the macromolecular composition. Using permeable cells of *E. coli* that have a deletion in the chromosomal *lpp* or *trpA* gene and are transformed with multicopy plasmids carrying the corresponding gene, polyadenylated *lpp* mRNA or *trpA* mRNA was synthesized with poly(A) tracts of about 14 nucleotides, similar in length to those found in the corresponding mRNAs isolated from growing cells (14, 15). Recently, permeable cells have been used to study the addition of poly(A) tracts to the 3′-ends of endogenous mRNA by inhibiting de novo RNA synthesis with rifampicin or streptolydigin (66). As described in the section on Relationships Between Poly(A) Polymerases and

PARSE EX.1082.013

3′-Exonucleases and Their Effect on mRNA Half-Life, the permeable cell system lends itself to the study of the dynamics of 3′-end processing of bacterial mRNA.

POSSIBLE FUNCTIONS OF RNA POLYADENYLATION IN BACTERIA

Even though eukaryotic poly(A) RNA has been under intensive investigation for 25 years, its function is still not entirely clear. Even less is known about bacterial poly(A) RNA, which began to be studied on the molecular level only 5 years ago (20, 45), and the discussion of the functions of bacterial RNA polyadenylation is therefore largely speculative. What is clear is that these functions are quite diverse, ranging from the control of plasmid replication to the modulation of mRNA stability and to a possible role in mRNA translation.

Modulation of the Control of Plasmid Regulation by Antisense RNA

The *pcnB* gene, which encodes the major poly(A) polymerase of *E. coli* (45), was originally identified as controlling the replication of ColE1 plasmids, *pcnB* mutants having much reduced copy numbers of plasmids such as pBR322 (46). Plasmid copy number control involves an antisense RNA (RNAI), which anneals to the RNA primer required for the initiation of plasmid replication (RNAII), thereby interfering with its function (for review see 67). The control of the intracellular concentration of RNAI by ribonucleases is thus a critical determinant of plasmid replication (68). Two laboratories (51, 69) independently made the seminal discovery that the product of the *pcnB* gene is required for the rapid degradation of RNAI so as to allow the initiation of plasmid replication by RNAII. The degradation of RNAI is initiated by the loss of a 5′-terminal pentanucleotide through the action of RNase E, yielding the cleavage product, $RNAI_{-5}$, which still represses plasmid replication but is destined for rapid degradation. The further degradation of $RNAI_{-5}$ to inactive molecules is controlled by the *pcnB* gene (51), and polyadenylation has been shown to play a key role in this process (51). Polyadenylated $RNAI_{-5}$ has the short half-life (1–2 min) typical of mRNA, whereas unadenylated $RNAI_{-5}$ is much more stable ($t_{1/2} > 10$ min). The accelerated degradation of polyadenylated $RNAI_{-5}$ is primarily caused by the 3′-exonucleolytic action of PNPase, which can initiate degradation only inefficiently adjacent to stem-loop structures such as those at the 3′-end of $RNAI_{-5}$ but, given a "toehold" on a 3′-terminal poly(A) tract, can proceed more rapidly through the stem loop, owing to its processivity (70, 71). Analysis of *pnp* mutants suggests that, in addition, PNPase may facilitate cleavage of RNAI by RNase E (70). Besides RNase E and PNPase, other nucleases

PARSE EX.1082.014

may participate in the inactivation of RNAI, including an enzyme that acts at the 5′-end of $RNAI_{-5}$ (70) and RNase III (M Masters, personal communication). Interestingly, mutations in the *rnc* gene, which lead to the loss of RNase III and result in a reduced rate of degradation of the polyadenylated forms of RNAI, have no effect on plasmid copy number, implying that polyadenylated RNAI is not an inhibitor of plasmid replication (M Masters, personal communication). According to this view, the polyadenylation of RNAI has two effects: It inactivates RNAI as an inhibitor of plasmid replication, perhaps by promoting a change in RNA secondary structure, and facilitates its degradation by 3′-exonucleases.

Relationships Between Poly(A) Polymerases and 3′-Exonucleases and Their Effect on mRNA Half-Life

Does polyadenylation destabilize only the antisense RNAs involved in plasmid replication or also host mRNA? There is some evidence that poly(A) tracts do indeed destabilize certain types of mRNA. This includes the observations that the half-lives of *lpp*, *trxA*, and *ompA* mRNA (72) and of *rpsO* mRNA (22) are significantly increased by deletion of *pcnB* when studied in a *pnp*, *rnb*, and *rne* background. In every case, mRNA stabilization was associated with reduced levels of polyadenylation. It should be noted that these results were obtained in the absence of RNase E and that the 3′-ends of most mRNA therefore consist of the stem-loop structures of rho-independent transcription terminators. Since exonucleolytic degradation of RNA by PNPase and RNase II (73–76) is strongly impeded by terminal stem-loop structures, polyadenylation of primary transcripts resulting from rho-independent transcription termination would provide a toehold for launching the processive exonucleolytic degradation by these enzymes (70, 71, 73, 74).

Most of the experiments on destabilization of mRNA by polyadenylation cited above were done in mutants lacking various ribonucleases, and it is not certain whether the conclusions drawn from them apply to normal mRNA turnover in wild-type *E. coli*. To address this question, I review the types of reactions that can occur at the 3′-ends of mRNAs, focusing on the most common transcripts, which result from rho-independent transcription termination. As illustrated in Figure 2, the primary transcript can either be polyadenylated at its 3′-end or degraded by 3′-exonucleases, although the latter process would be very slow, owing to the terminal stem-loop structure (73, 74, 76). The polyadenylated transcript, no longer terminated by a stem loop, is more susceptible than the primary transcript to 3′-exonucleolytic degradation by enzymes such as RNase II and polynucleotide phosphorylase. Another possible fate of the primary transcript (as well as its polyadenylation product) is cleavage at a specific site at the base of the terminal stem-loop structure by RNase E (77), an enzyme that has been

PARSE EX.1082.015

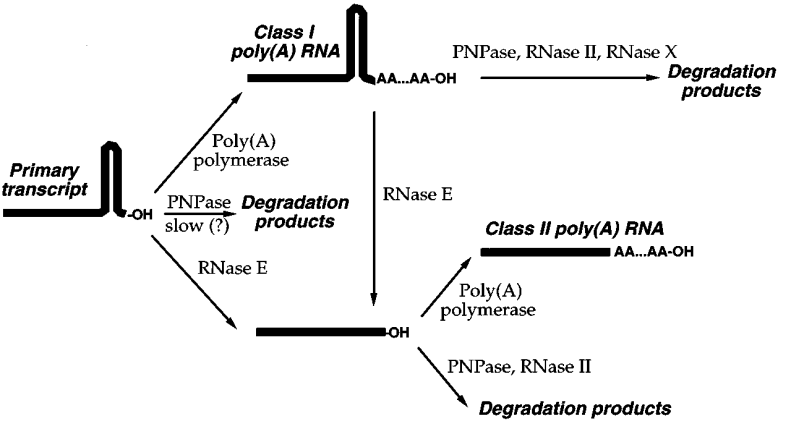

Figure 2 Postulated steps in the processing of primary transcripts produced by rho-independent transcription termination, with particular focus on the roles of RNase E, PNPase, RNase II, and poly(A) polymerases. See text for discussion. RNase *X* refers to an unknown 3′-exoribonuclease, postulated to account for the observation that mRNA degradation occurs efficiently even in the absence of RNase E, PNPase, and RNase II (22, 72).

implicated in the control of mRNA stability (78–80). The cleaved transcript, no longer protected from 3′-exonucleolytic degradation by the terminal secondary structure, would be highly vulnerable to degradation by 3′-exonucleases such as polynucleotide phosphorylase or RNase II in competition with polyadenylation. This explains what might seem to be a paradox: Under certain conditions, i.e. when mRNA has a 3′-terminal stem-loop structure, polyadenylation acts synergistically with 3′-exonucleases to promote RNA degradation; under other conditions, i.e. when the 3′-termini of mRNA are not stabilized by secondary structure, polyadenylation competes with 3′-exonucleases and promotes mRNA stabilization. The former condition, which results in mRNA destabilization by polyadenylation, would apply to most mRNA in *E. coli rne* strains, which lack RNase E and thus tend to retain the stem loops of the rho-independent terminator, as illustrated in the top part of Figure 2 and supported by the high ratio of Class I to Class II poly(A) RNA (see Figure 1) in such strains (21, 22). Indeed, many of the studies on the destabilization of mRNA by polyadenylation cited above were carried out in *rne* mutants. The latter condition, in which polyadenylation and exonucleolytic degradation are competitive processes, would be observed in rne^+ strains, where the reactions in the bottom half of Figure 2 play a significant role, as indicated by the high ratio of Class II to Class I poly(A) RNA (see Figure 1) in such strains (20–22). The observation that rne^+ strains of *E. coli* lacking the two major 3′-exonucleases have higher cellular

PARSE EX.1082.016

concentrations of polyadenylated mRNA (20) is evidence for this type of competition. Under such conditions, the control of mRNA half-life would be highly dependent on the ratio of 3′-exonuclease to poly(A) polymerase activities. This prediction is supported by measurements of mRNA polyadenylation in permeable cells under conditions in which 3′-exoribonucleases are inactive, which showed a sixfold higher rate of poly(A) synthesis in exonuclease-deficient mutant strains, suggesting a higher concentration of mRNA 3′-ends amenable to polyadenylation (66). These results imply that 3′-exonucleases reduce the level of polyadenylated mRNA in *E. coli* not merely by causing its degradation but also by reducing its rate of synthesis, presumably by competing with poly(A) polymerase for the 3′-ends of mRNA.

However, the picture of RNA degradation at the 3′-ends of mRNA may be considerably more complex than that shown in Figure 2, which assumes that endonucleases, exonucleases, and poly(A) polymerases act independently. In the past two years, evidence has accumulated that the exonucleolytic degradation and endonucleolytic cleavage of *E. coli* mRNA may be coordinated by a multienzyme complex termed the degradosome. A complex of polynucleotide phosphorylase and RNase E, capable of processing the 3′-ends of RNA, it was first described by Carpousis et al (81) and was subsequently found to contain Rh1B, a protein of the DEAD-box family of ATP-dependent RNA helicases; DnaK, a heat-shock protein; and enolase, a glycolytic enzyme with no known function in RNA metabolism (82, 83). The complete degradation of RNA with a 3′-terminal stem-loop structure by the degradosome complex requires the presence of ATP and is inhibited by anti-Rh1B antibody, suggesting that RNA helicase activity facilitates the processive degradation by polynucleotide phosphorylase of RNAs with a high degree of secondary structure (83). An alternative proposal, based on the observation that the DEAD-box proteins protect RNA from endonucleases (84), is that Rh1B protein modulates RNase E activity (82). The degradosome may also participate in the degradation of antisense RNAI, which is essential for ColE1 plasmid replication and is activated by 3′-polyadenylation as discussed above. Attempts to detect PAP I, the product of the *pcnB* gene, in degradosomes were unsuccessful (83), but the possibility that PAP II (53) may be associated with degradosomes has not yet been tested.

Some progress has been made in establishing in vitro systems for examining the relationship between the polyadenylation and degradation of *E. coli* mRNA. Studies with ribosomal protein S20 mRNA showed that polyadenylation of the 3′-terminal stem-loop structure of rho-independent transcription terminators allows complete degradation by polynucleotide phosphorylase, whereas RNase II cannot degrade the intact primary transcript even if it is polyadenylated, but can initiate degradation of 3′-polyadenylated RNase E cleavage products in regions protected by relatively weak secondary structure (74). Another promising

PARSE EX.1082.017

in vitro system, which approximates the in vivo decay rates for *lpp* and *trxA* mRNA, consists of purified polysomes, with which PAP I and ribonucleases appear to be tightly associated (85). In the absence of ATP, *lpp* and *trxA* mRNA are degraded slowly, but their decay rates are greatly enhanced by the addition of ATP, approaching those observed in vivo. This response to ATP is not observed when polysomes are derived from *pcnB* mutant strains, suggesting that it involves the PAP I-mediated synthesis of poly(A) tracts (85). These observations lend strong support to the notion that mRNA polyadenylation plays an important role in the initiation of exonucleolytic degradation.

Involvement in Translation of mRNA

The poly(A) tracts of eukaryotic mRNA are invariably associated with poly(A) binding proteins (PABP). A 70-kDa PABP has been characterized from yeast and found to be essential for growth (86, 87). Considerable evidence shows that PABP plays a role in the regulation both of mRNA stability and of mRNA translation (for a review see 88).

The search for an analogous PABP in *E. coli* led to the identification and purification of a 61-kDa protein that was found to be identical to the small ribosomal subunit protein S1 (MP Kalapos & N Sarkar, submitted for publication). S1 binds to poly(A) in a mildly cooperative manner with an association constant of 3×10^6 M^{-1}. When poly(A) tracts, together with any bound materials, were isolated from polysomes by ribonuclease digestion and oligo(dT) cellulose chromatography and subjected to Western blot analysis with S1-specific antibodies, S1 protein was detected, suggesting an in vivo association of S1 with the poly(A) tracts of *E. coli* mRNA (MP Kalapos & Sarkar, submitted for publication). The functional significance of the binding of S1 protein to poly(A) RNA is still uncertain, but modulation of the efficiency of translation is an interesting possibility. In yeast, a translational role for poly(A) and PABP is supported by the observations that polyadenylation stimulates in vitro mRNA translation (89–91) and that the complex of PABP with poly(A) tracts helps in recruiting the 40S ribosomal subunit to mRNA (92). The latter effect is reminiscent of the role of S1 protein in mRNA recruitment to the *E coli* 30S ribosome (93) and suggests that binding to poly(A) tracts may play a role in the stimulation of translation initiation by S1 protein.

mRNA POLYADENYLATION IN EUKARYOTIC ORGANELLES

The notion that eukaryotic organelles such as mitochondria and chloroplasts are derived from endosymbiotic prokaryotes is now generally accepted (94). It is thus of interest to review the process of RNA polyadenylation in these organelles

PARSE EX.1082.018

to determine whether it reflects the prokaryotic or the eukaryotic pattern. This brief summary addresses only one aspect of organellar mRNA polyadenylation and focuses only on mammalian mitochondria and chloroplasts from higher plants.

Mitochondria

A hallmark of the animal mitochondrial genome is its extreme economy: More than 90% of the genome encodes either proteins or RNA components of the translational machinery (95). Transcription starts bidirectionally at a unique region of the circular mitochondrial genome and yields two polycistronic RNA molecules in which protein coding regions are interspersed with transfer RNA genes. These RNA molecules are processed by endonucleolytic cleavage adjacent to the tRNA genes to yield mRNA molecules, each of which encodes a single protein (96). An interesting feature of mitochondrial mRNA is that there are no 5′ and 3′ untranslated regions; the mRNA starts at the AUG translation initiation codon and ends with the U or the UA of the termination codon, which is converted to a functional translation stop signal by polyadenylation (96). The sequence of events in the processing of the primary mitochondrial transcript to functional mRNA—endonucleolytic cleavage followed by polyadenylation—is not unlike the sequence of processing events in the *E. coli lacZY* intercistronic region discussed above (30), and mitochondrial mRNAs are thus analogous to bacterial Class IV poly(A) RNA (see Figure 1). The fact that the major function of the polyadenylation process itself is to complete the translation termination codon in many mRNAs suggests that it occurs indiscriminately at all RNA ends in the absence of specific recognition sequences, and it therefore resembles bacterial mRNA polyadenylation rather than the eukaryotic process. The transient polyadenylation of mitochondrial ribosomal RNA (96–99) may be a consequence of indiscriminate mitochondrial RNA polyadenylation. An alternate view is that polyadenylation is a highly specific process that is tightly linked to the endonucleolytic cleavage step that releases the 3′-ends of the RNA products from the flanking tRNA sequences in mitochondria (96). The average size of the poly(A) tracts associated with mitochondrial RNA from HeLa cells is about 55 nucleotides (100), and poly(A) tracts from pulse-labeled mRNAs in mouse ascites cell mitochondria range between 35 and 55 adenylate residues (101), comparable in size to bacterial poly(A) tracts.

The poly(A) polymerase from Morris hepatoma cell mitochondria has a molecular weight of 60,000 and can add poly(A) tracts with an average length of 600 nucleotides to mitochondrial RNA primers (102). Studies on the synthesis of poly(A) RNA in isolated mitochondria demonstrated the posttranscriptional addition of 20–23 adenylate residues (103). A coding sequence for poly(A) polymerase has not been identified in the mammalian mitochondrial

PARSE EX.1082.019

genome, implying that mitochondrial poly(A) polymerase is encoded by a nuclear gene.

Chloroplasts

Although the polyadenylation of plant chloroplast mRNA was described in 1976 (104), this process was not further investigated until very recently (37, 105). Recent studies reveal some interesting structural differences as well as functional similarities with mRNA polyadenylation in *E. coli*.

Chloroplast poly(A) tracts are several hundred nucleotides in length, in the same range as eukaryotic poly(A) RNA but much longer than bacterial poly(A) tracts; but unlike both eukaryotic and bacterial poly(A) tracts, they are not necessarily adenylate homopolymers but, in the case of *pbsA* mRNA, consist of adenylate clusters (75%) interspersed with guanylate (25%) and some cytidylate and uridylate residues (5%) (37). This phenomenon is reminiscent of the noncoded poly(A)-rich segments of bacteriophage T7 mRNA (36). Most polyadenylation sites in chloroplast *psbA* mRNA, analyzed by cDNA cloning, are within the truncated *pbsA* coding region, rather than at the 3′-end of the mature transcript (37), and correspond to Class V of bacterial poly(A) RNA (Figure 1).

A poly(A) polymerase, which has been purified from *Pisum sativum* and several other plants, is a complex consisting of PAP-I, a 43-kDa polypeptide immunologically related to the yeast poly(A) polymerase, and two nucleoprotein fractions, PAP-II and PAP-III, of which only PAP-III, a 105-kDa RNA binding protein, appears to be essential for the function of PAP-I, with its endogenous RNA serving as a primer for polyadenylation (106). Further analysis of PAP-III showed it to be a chloroplast-localized polynucleotide phosphorylase (PNPase), implying that PAP-I, with which PAP-III interacts, is a chloroplast-specific poly(A) polymerase (105). The association of poly(A) polymerase and PNPase with the 3′-ends of mRNA suggests that the two enzymes may act synergistically to mediate mRNA degradation, in a manner analogous to the events at the terminal stem-loop structures of bacterial mRNA that are illustrated in the top part of Figure 2. The primary transcripts from chloroplasts also have 3′-terminal stem-loop structures that serve as barriers to exonucleolytic digestion but can be removed by endonucleolytic cleavage (107, 108). A high-molecular-weight nucleoprotein complex has been identified that contains RNA binding proteins as well as endonuclease and exonuclease activities that appear to play a role in the degradation of polyadenylated 3′-ends of mRNA (37, 109). The low levels of Class I poly(A) RNA (see Figure 1) in chloroplasts compared to partially degraded mRNA molecules with polyadenylated truncated coding regions are consistent with this notion and support the idea that polyadenylation

PARSE EX.1082.020

plays a role in chloroplast mRNA degradation (37), by a mechanism analogous to that proposed in bacteria (see above).

Many poly(A)-binding proteins have been cloned in plants (110, 111), but none has yet been specificially identified with chloroplasts. However, the chloroplast small ribosomal protein, CS1, which has significant homology with *E. coli* S1 protein, has specific poly(A)-binding activity and may be involved in the formation of translation initiation complexes with chloroplast mRNA (112), analogous to the proposed role of bacterial S1 protein (MP Kalapos & N Sarkar, submitted for publication).

mRNA POLYADENYLATION IN ARCHAEA

Archaeal mRNA resembles that from eubacteria not only in its relative short half-life but in that it is associated with short 3′-terminal poly(A) tracts, which range from 10 nucleotides in *Methanococcus vannielii* (113) to 12–40 nucleotides in *Halobacterium halobium* (114). Analysis of an oligo(dT)-primed cDNA library from the hyperthermophile *Thermotoga maritima* revealed no polyadenylation at 3′-terminal stem-loop structures but many clones with poly(A) tracts just downstream of the coding region (115).

CONCLUDING REMARKS

The data summarized in this review leave no doubt that polyadenylation plays a major role in mRNA processing and function in both prokaryotes and eukaryotes. Nevertheless, comparison of prokaryotic and eukaryotic mRNA polyadenylation reveals important differences. Table 1 summarizes the characteristics of prokaryotic and eukaryotic mRNA polyadenylation, using *E. coli* and mammals as representative examples, with the caveat that there are also minor differences in the polyadenylation process between different bacterial species as well as between yeast and mammals. Most striking are the differences between *E. coli* and higher eukaryotes in the control of the specificity of polyadenylation and some aspects of its function.

The realization that mRNA polyadenylation is universal immediately suggests the notion that we are dealing with an ancient process that has originated in a common ancestor. On the other hand, in view of the differences summarized in Table 1 and the fact that the bacterial poly(A) polymerases show no sequence homology with their eukaryotic counterparts, we might draw a different conclusion. To resolve this apparent contradiction, we must attempt to deduce the original function of RNA polyadenylation in the universal ancestor and investigate whether all the diverse functions seen in bacteria and mammals today

PARSE EX.1082.021

Table 1 Comparison of mRNA polyadenylation in prokaryotes and eukaryotes

	Escherichia coli	Mammals
Length of poly(A) tracts	14–60 nucleotides	80–200 nucleotides
Extent of polyadenylation of any single mRNA species	2–50%	100%
Polyadenylation sites	Indiscriminate at accessible 3′-ends	Downstream of AAUAAA consensus element
Poly(A) polymerases	Multiple enzymes encoded by unrelated genes	Multiple enzyme forms encoded by a single gene
Functions		
mRNA stability	Destabilization of mRNA with terminal stem-loop structures; possible stabilization of mRNA without terminal stem-loops	Stabilization of mRNA without terminal stem-loop structures
Translation	Possible role through binding of S1 ribosomal protein to 30S ribosomes	Interaction with PABP may be involved in binding mRNA to 40S ribosomes
Plasmid replication	Degradation of inhibitory antisense RNA	

could have arisen by divergence. An important guiding principle in this kind of discussion is that most regulatory processes have evolved opportunistically and relatively recently from processes that at one time played (or still play) a role in the machinery for assembling biological macromolecules. We must therefore search for evolutionary antecedents by looking for analogous reactions with biosynthetic functions. In the case of RNA polyadenylation, such a prototype might be the posttranscriptional addition of a terminal adenylate residue to transfer RNA to make it a functional amino acid acceptor. The discoveries of significant sequence homology between the *E. coli* tRNA nucleotidyltransferase and the major *E. coli* poly(A) polymerase by Masters et al (49), and between the tRNA nucleotidyltransferase of the archeon *Sulfolobus shibatae* and eukaryotic poly(A) polymerases (116), support this idea. A vestige of this ancestral function may still be seen in mRNA polyadenylation in mammalian mitochondria, which serves to provide mRNA with functional translation termination codons (96). It is not difficult to visualize how, from this ancestral role in macromolecular synthesis, polyadenylation could have evolved to assume regulatory functions as diverse as controlling the stability of mRNA as well as of the regulatory RNAs that modulate DNA replication and serving as a site for the binding of regulatory proteins to the 3′-ends of mRNA with the capacity to modulate its binding to ribosomes or its stability. It is this divergence from

PARSE EX.1082.022

a biosynthetic function into the regulatory arena, driven by selective pressures imposed by the physiological needs of different organisms, that could well have given rise to the differences summarized in Table 1.

ACKNOWLEDGMENTS

Research in the author's laboratory is supported by NIH grant GM26517 and by a Public Health Service Fogarty International Research Fellowship TW04657 to Dr. MP Kalapos. I thank the colleagues who sent preprints and manuscripts prior to publication. Special thanks are due to the past and present members of my laboratory for their contributions to the study of bacterial poly(A) RNA and to Dorothy Syrigos for skillful preparation of the manuscript. Finally, I thank Dr. Henry Paulus to whom I am indebted for assistance with the preparation of this review, stimulating discussions, and support throughout the many years of research on bacterial poly(A) RNA when it was viewed by most others as merely a scientific curiosity.

Visit the *Annual Reviews home page* at http://www.annurev.org.

Literature Cited

1. Sarkar N. 1996. *Microbiology* 142:3125–33
2. August JT, Ortiz PJ, Hurwitz J. 1962. *J. Biol. Chem.* 237:3786–93
3. Nakazato H, Venkatesan S, Edmonds M. 1975. *Nature* 256:144–46
4. Srinivasan PR, Ramanarayanan M, Rabbani E. 1975. *Proc. Natl. Acad. Sci. USA* 72:2910–14
5. Ohta N, Sanders M, Newton A. 1975. *Proc. Natl. Acad. Sci. USA* 72:2343–46
6. Sarkar N, Langley D, Paulus H. 1978. *Biochemistry* 17:3468–74
7. Schultz GA, Chaconas G, Moore RL. 1978. *J. Bacteriol.* 133:569–75
8. Gopalakrishna Y, Langley D, Sarkar N. 1981. *Nucleic Acids Res.* 9:3545–54
9. Gopalakrishna Y, Sarkar N. 1982. *Biochemistry* 21:2724–29
10. Crouch DH, Ownby JD, Carr NG. 1983. *J. Bacteriol.* 156:979–82
11. Käufer N, Altmann M, Döhren HV. 1981. *FEMS Microbiol. Lett.* 12:71–75
12. Yu P-L, Drews G. 1982. *FEMS Microbiol. Lett.* 14:233–36
13. Majumdar PK, McFadden BA. 1984. *J. Bacteriol.* 157:795–801
14. Taljanidisz J, Karnik P, Sarkar N. 1987. *J. Mol. Biol.* 193:507–15
15. Karnik P, Taljanidisz J, Sasvari-Szekely M, Sarkar N. 1987. *J. Mol. Biol.* 196:347–54
16. Gopalakrishna Y, Sarkar N. 1982. *J. Biol. Chem.* 257:2747–50
17. Karnik P, Gopalakrishna Y, Sarkar N. 1986. *Gene* 49:161–65
18. Kushner SR. 1996. See Ref. 117, pp. 849–60
19. Nakamura K, Pirtle RM, Pirtle IM, Takeishi K, Inouye M. 1980. *J. Biol. Chem.* 255:210–16
20. Cao G-J, Sarkar N. 1992. *Proc. Natl. Acad. Sci. USA* 89:7546–50
21. Haugel-Nielsen J, Hajnsdorf E, Régnier P. 1996. *EMBO J.* 15:3144–52
22. Hajnsdorf E, Braun F, Haugel-Nielsen J, Régnier P. 1995. *Proc. Natl. Acad. Sci. USA* 92:3973–77
23. Aiba H, Fujimoto S, Ozaki N. 1982. *Nucleic Acids Res.* 10:1345–61
24. Cossart P, Gicquel-Sanzey B. 1982. *Nucleic Acids Res.* 10:1363–78
25. Alifano P, Rivellini F, Limauro D, Bruni CB, Carlomagno MS. 1991. *Cell* 64:553–63
26. Wu AM, Christie GE, Platt T. 1981. *Proc. Natl. Acad. Sci. USA* 78:2913–17
27. Sarkar N, Cao G-J, Kalapos MP. 1996.

PARSE EX.1082.023

Cold Spring Harbor Meet., Cold Spring Harbor, NY (Abstr.)
28. Sarkar N, Cao G-J. 1996. *1996 FASEB Summer Conf.*, Copper Mountain, CO (Abstr.)
29. Murakawa GJ, Kwan C, Yamashita J, Nierlich DP. 1991. *J. Bacteriol.* 173:28–36
30. Kwan C, Nierlich DP. 1996. *Am. Soc. Microbiol. Meet.*, New Orleans (Abstr.)
31. Matsumoto Y, Shigesada K, Hirano M, Imai M. 1986. *J. Bacteriol.* 166:945–58
32. Landick R, Turnbough J, CL, Yanofsky C. 1996. See Ref. 117, pp. 1263–86
33. Cao G-J, Sarkar N. 1993. *FEMS Microbiol. Lett.* 108:281–86
34. Dunn JJ, Studier FW. 1983. *J. Mol. Biol.* 166:477–535
35. Kramer RA, Rosenberg M, Steitz JA. 1974. *J. Mol. Biol.* 89:767–76
36. Johnson MD. 1989. *Polyadenylation of bacteriophage T7 mRNA.* Undergraduate honors thesis. Harvard Univ., Cambridge
37. Lisitsky I, Klaff P, Schuster G. 1996. *Proc. Natl. Acad. Sci. USA* 93:13398–403
38. Rosenberg M, Weissman S, deCrombrugghe B. 1975. *J. Biol. Chem.* 250:4755–64
39. Smith GR, Hedgpeth J. 1975. *J. Biol. Chem.* 250:4818–21
40. Proudfoot N. 1991. *Cell* 64:671–74
41. Edmonds M, Abrams R. 1960. *J. Biol. Chem.* 235:1142–49
42. Schäfer R, Zillig W, Priess H. 1972. *FEBS Lett.* 25:87–90
43. Sippel AE. 1973. *Eur. J. Biochem.* 37:31–40
44. Ramanarayanan M, Srinivasan PR. 1976. *J. Biol. Chem.* 251:6274–86
45. Cao G-J, Sarkar N. 1992. *Proc. Natl. Acad. Sci. USA* 89:10380–84
46. Lopilato J, Bortner S, Beckwith J. 1986. *Mol. Gen. Genet.* 205:285–90
47. Liu J, Parkinson JS. 1989. *J. Bacteriol.* 171:1254–61
48. Régnier P, Grunberg-Manago M, Portier C. 1987. *J. Biol. Chem.* 262:63–68
49. Masters M, March JB, Oliver IR, Collins JF. 1990. *Mol. Gen. Genet.* 220:341–44
50. Masters M, Colloms MD, Oliver IR, He L, MacNaughton EJ, Charters Y. 1993. *J. Bacteriol.* 175:4405–13
51. Xu F-F, Lin-Chao S, Cohen SN. 1993. *Proc. Natl. Acad. Sci. USA* 90:6756–60
52. Kalapos MP, Cao G-J, Kushner SR, Sarkar N. 1994. *Biochem. Biophys. Res. Commun.* 198:459–65
53. Cao G-J, Pogliano J, Sarkar N. 1996. *Proc. Natl. Acad. Sci. USA* 93:11580–85
54. Donovan WP, Kushner SR. 1986. *Proc. Natl. Acad. Sci. USA* 83:120–24
55. Reuven NB, Deutscher MP. 1993. *FASEB J.* 7:143–48
56. Itaya M. 1990. *Proc. Natl. Acad. Sci. USA* 87:8587–91
57. Ryner LC, Takagaki Y, Manley JL. 1989. *Mol. Cell. Biol.* 9:4229–38
58. Thuresson A-C, Åström J, Åström A, Grönvik K-O, Virtanen A. 1994. *Proc. Natl. Acad. Sci. USA* 91:979–83
59. Lingner J, Kellermann J, Keller W. 1991. *Nature* 354:496–98
60. Sorokin AV, Azevedo V, Zumstein E, Galleron N, Ehrlich SD, Serror P. 1995. *GenBank* Access. L47709
61. DeRossi E, Cantoni R, Labo M, Riccardo G. 1995. *GenBank* Access. L39923
62. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. 1995. *Science* 269:496–512
63. Nakamura Y, Kaneko T, Tabata S, Miyajima N. 1996. *www.kazusa.or.jp/cyano.html.* Sequences sll0825 and sll1253.
64. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, et al. 1995. *Science* 270:397–403
65. Gopalakrishna Y, Sarkar N. 1983. *Arch. Biochem. Biophys.* 224:196–205
66. Cao G-J, Kalapos MP, Sarkar N. 1997. *Biochimie*. In press
67. Polisky B. 1988. *Cell* 55:929–32
68. Lin-Chao S, Cohen SN. 1991. *Cell* 65:1233–42
69. He L, Söderbom F, Wagner EGH, Binnie U, Binns N, Masters M. 1993. *Mol. Microbiol.* 9:1131–42
70. Xu F-F, Cohen SN. 1995. *Nature* 374:180–83
71. Cohen SN. 1995. *Cell* 80:829–32
72. O'Hara EB, Chekanova JA, Ingle CA, Kushner ZR, Peters E, Kushner SR. 1995. *Proc. Natl. Acad. Sci. USA* 92:1807–11
73. Coburn GA, Mackie GA. 1996. *J. Biol. Chem.* 271:1048–53
74. Coburn GA, Mackie GA. 1996. *J. Biol. Chem.* 271:15776–81
75. Guarneros G, Portier C. 1990. *Biochimie* 72:771–77
76. McLaren RS, Newbury SF, Dance GSC, Causton HC, Higgins CF. 1991. *J. Mol. Biol.* 221:81–95
77. Régnier P, Hajnsdorf E. 1991. *J. Mol. Biol.* 217:283–92
78. Ono M, Kuwano M. 1979. *J. Mol. Biol.* 129:343–57
79. Arraiano CM, Yancey SD, Kushner SR. 1988. *J. Bacteriol.* 170:4625–33
80. Ehretsmann CP, Carpousis AJ, Krisch HM. 1992. *FASEB J.* 6:3186–92
81. Carpousis AJ, Van Houwe G, Ehretsmann C, Krisch HM. 1994. *Cell* 76:889–900

PARSE EX.1082.024

82. Miczak A, Kaberdin VR, Wei C-L, Lin-Chao S. 1996. *Proc. Natl. Acad. Sci. USA* 93:3865–69
83. Py B, Higgins CF, Krisch HM, Carpousis AJ. 1996. *Nature* 381:169–72
84. Iost I, Dreyfus M. 1994. *Nature* 372:193–96
85. Ingle CA, Kushner SR. 1996. *Proc. Natl. Acad. Sci. USA* 93:12926–31
86. Sachs AB, Bond MW, Kornberg RD. 1986. *Cell* 45:827–35
87. Sachs A. 1990. *Curr. Opin. Cell Biol.* 2: 1092–98
88. Jacobson A. 1996. In *Translation Control*, ed. JWB Hershey, MB Mathews, N Sonenberg, pp. 451–80. Plainview, NY: Cold Spring Harbor Lab. Press
89. Munroe D, Jacobson A. 1990. *Mol. Cell. Biol.* 10:3441–55
90. Jackson RJ, Standart N. 1990. *Cell* 62:15–24
91. Sheets MD, Fox CA, Hunt T, Van de Woude G, Wickens M. 1994. *Genes Dev.* 8:926–38
92. Tarun SZ Jr., Sachs AB. 1995. *Genes Dev.* 9:2997–3007
93. Subramanian AR. 1983. *Progr. Nucleic Acid Res. Mol. Biol.* 28:101–43
94. Gray MW. 1992. *Int. Rev. Cytol.* 141:233–357
95. Attardi G. 1985. *Int. Rev. Cytol.* 93:93–145
96. Ojala D, Montoya J, Attardi G. 1981. *Nature* 290:470–74
97. Dubin DT, Timko KD, Baer RJ. 1981. *Cell* 23:271–78
98. Berthier F, Renaud M, Alziari S, Durand R. 1986. *Nucleic Acids Res.* 14:4519–33
99. Benkel BF, Duschesnay P, Boer PH, Genest Y, Hickey DA. 1988. *Nucleic Acids Res.* 16:9880
100. Hirsch M, Penman S. 1973. *J. Mol. Biol.* 80:379–91
101. Avadhani NG. 1979. *Biochemistry* 18: 2673–78
102. Rose KM, Morris HP, Jacob ST. 1975. *Biochemistry* 14:1025–32
103. Rose KM, Jacob ST. 1976. *Biochemistry* 15:5046–52
104. Haff LA, Bogorad L. 1976. *Biochemistry* 15:4110–15
105. Li Q-S, Das Gupta J, Hunt AG. 1996. *J. Biol. Chem.* 271:19831–35
106. Das Gupta J, Li Q-S, Thomson AB, Hunt AG. 1995. *Plant Sci.* 110:215–26
107. Stern DB, Kindle KL. 1993. *Mol. Cell. Biol.* 13:2277–85
108. Drager RG, Zeidler M, Simpson CL, Stern DB. 1996. *RNA* 2:652–63
109. Hayes R, Kudla J, Schuster G, Gabay L, Maliga P, Gruissem W. 1996. *EMBO J.* 15:1132–41
110. Belostotsky DA, Meagher RB. 1993. *Proc. Natl. Acad. Sci. USA* 90:6686–90
111. Yang J-J, Hunt AG. 1992. *Plant Physiol.* 98:1115–20
112. Franzetti B, Carol P, Mache R. 1992. *J. Biol. Chem.* 267:19075–81
113. Brown JW, Reeve JN. 1985. *J. Bacteriol.* 162:909–17
114. Brown JW, Reeve JN. 1986. *J. Bacteriol.* 166:686–88
115. Kim CW, Markiewicz P, Lee JL, Schierle CF, Miller JH. 1993. *J. Mol. Biol.* 231:960–81
116. Yu D, Maizels N, Weiner A. 1996. *RNA* 2:895–908
117. Neidhardt FC, ed. 1996. Escherichia coli *and* Salmonella: *Cellular and Molecular Biology*, Vol. 1. Washington, DC: ASM

PARSE EX.1082.025