Acta Biochimica Polonica

Vol. 45 No. 4/1998

895-906

QUARTERLY

This paper is dedicated to Professor Maciej Wiewiórowski

Review

Cyclases of the 3′-terminal phosphate in RNA: A new family of RNA processing enzymes conserved in Eucarya, Bacteria and Archaea

Witold Filipowicz✉, Eric Billy, Krzysztof Drabikowski and Pascal Genschik★

Friedrich Miescher-Institut, 4002 Basel, Switzerland

Received: 22 September, 1998

Key words: adenylylation, cyclic phosphate, RNA ligation, RNA modification, RNA processing

The 2′,3′-cyclic phosphate termini are produced, as either intermediates or final products, during RNA cleavage by many different endoribonucleases. Likewise, ribozymes such as hammerheads, hairpins, or the hepatitis delta ribozyme, generate 2′,3′-cyclic phosphate ends. Discovery of the RNA 3′-terminal phosphate cyclase has indicated that cyclic phosphate termini in RNA can also be produced by an entirely different mechanism. The RNA 3′-phosphate cyclase converts the 3′-terminal phosphate in RNA into the 2′,3′-cyclic phosphodiester in the ATP-dependent reaction which involves formation of the covalent cyclase-AMP and the RNA-$N^{3'}$ $pp^{5'}$A intermediates. The findings that several eukaryotic and prokaryotic RNA ligases require the 2′,3′-cyclic phosphate for the ligation of RNA molecules raised a possibility that the RNA 3′-phosphate cyclase may have an anabolic function in RNA metabolism by generating terminal cyclic groups required for ligation. Recent cloning of a cDNA encoding the human cyclase indicated that genes encoding cyclase-like proteins are conserved among Eucarya, Bacteria, and Archaea. The protein encoded by the

✉Corresponding author: W. Filipowicz; Friedrich Miescher-Institut, P.O. Box 2543, 4002 Basel, Switzerland; Address for couriers: Maulbeerstrasse 66, 4058 Basel, Switzerland; tel: (41 61) 697 4128 or 697 6993; fax: (41 61) 697 3976; e-mail: Filipowi@fmi.ch

★Present address: Institut de Biologie Moléculaire des Plantes du C.N.R.S., 12 rue du Général Zimmer, 67084 Strasbourg, France

Abbreviations: AMPCPP, α,β-methylene analog of ADP; AMPPCP, β,γ-methylene analog of ATP; AMPPNP, β,γ-imido analog of ATP; EST, expressed sequence tags.

***Escherichia coli* gene was overexpressed and shown to have the RNA 3′-phosphate cyclase activity. This article reviews properties of the human and bacterial cyclases, their mechanism of action and substrate specificity. Possible biological functions of the enzymes are also discussed.**

The 2′,3′-cyclic phosphate termini are produced during RNA cleavage by different endoribonucleases. For many secretory degradative nucleases, such as RNases A or T1, cyclic phosphates are formed as intermediates which are subsequently opened into 3′-monoesters. However, for tRNA-splicing endonucleases from Eucarya and Archaea, the cyclic phosphate is a final product of the cleavage (reviewed by Phizicky & Greer, 1993; Westaway & Abelson, 1995). In addition, the type I topoisomerase has recently been shown to have endoribonuclease activity yielding 2′,3′-cyclic phosphate termini (Segikuchi & Shuman, 1997). Ribozymes such as hammerheads, hairpins, or the hepatitis delta ribozyme, also generate 2′,3′-cyclic phosphate ends (reviewed by Symons, 1992).

The fact that the 2′,3′-cyclic phosphate has an anabolic function in RNA metabolism emerged when it was found that eukaryotic RNA ligases require 2′,3′-cyclic ends for RNA ligation (Konarska *et al.*, 1981; 1982; Filipowicz *et al.*, 1983; Filipowicz & Shatkin, 1983; Furneaux *et al.*, 1983; Greer *et al.*, 1983a; Schwartz *et al.*, 1983; reviewed by Filipowicz & Gross, 1984; Phizicky & Greer, 1993; Westaway & Abelson, 1995). This requirement applies to each of the two RNA ligases known to be involved in nuclear pre-tRNA splicing in eukaryotes. One of the ligases, an enzyme generating the 3′,5′-phosphodiester, 2′-phosphomonoester linkage, functions additionally in splicing of the unusual intron present in *HAC1* pre-mRNA in yeast (Sidrauski & Walter, 1997). This ligase may also be involved in ligation of virusoid and viroid RNAs in plants (reviewed by Genschik *et al.*, 1998). Notably, the only known cellular RNA ligase in eubacteria, which joins RNA ends *via* the 2′,5′-phosphodiester, also requires 2′,3′-cyclic ends for ligation (Greer *et al.*, 1983b; Arn & Abelson, 1996).

In light of the reported role of cyclic phosphate termini in RNA metabolism, it was interesting to discover that endonucleolytic cleavage is not the only way to generate RNA molecules bearing 2′,3′-cyclic phosphates. Such molecules can also be produced by the action of the RNA 3′-terminal phosphate cyclase, an enzyme that catalyzes ATP-dependent conversion of a 3′-phosphate at the end of RNA to the 2′,3′-cyclic phosphodiester. RNA 3′-terminal phosphate cyclase has originally been identified in extracts of HeLa cells and *Xenopus* nuclei (Filipowicz *et al.*, 1983; Filipowicz & Shatkin, 1983). The cyclization of the 3′-terminal phosphate, catalyzed by the enzyme, occurs in three steps (Filipowicz *et al.*, 1985; Reinberg *et al.*, 1985; Vicente & Filipowicz, 1988; reviewed by Filipowicz & Vicente, 1990):

(a) Enzyme + ATP → Enzyme-AMP + PP_i

(b) RNA-$N^{3'}$p + Enzyme-AMP → RNA-$N^{3'}pp^{5'}$A + Enzyme

(c) RNA-$N^{3'}pp^{5'}$A → RNA-N > p + AMP

Support for step (a) comes from identification of the covalent cyclase-AMP complex and the ability of 3′-phosphorylated RNA but not the 3′-OH-terminated RNA to release AMP from the preformed cyclase-AMP complex (Filipowicz *et al.*, 1985; Reinberg *et al.*, 1985; Vicente & Filipowicz, 1988; Genschik *et al.*, 1997; 1998). Step (b) is inferred from experiments demonstrating accumulation of the RNA-$N^{3'}pp^{5'}$A molecules when the ribose at the RNA 3′-terminus is replaced with the 2′-deoxy- or 2′-*O*-methylribose (Filipowicz *et al.*, 1985). Reaction (c) probably occurs nonenzy-

matically as the result of nucleophilic attack by the adjacent 2′-OH on the phosphorus in the phosphodiester linkage. Mechanistically, with respect to covalent modification of the protein and a transfer of the nucleotidyl group from the protein-NMP intermediate to the terminal phosphate or pyrophosphate in nucleic acids, the cyclase resembles RNA and DNA ligases, and capping enzymes (reviewed by Shuman & Schwer, 1995). However, recent characterisation of the genes encoding the RNA 3′-terminal phosphate cyclase in humans and in *Escherichia coli* has indicated no structural relationship between these enzymes and the cyclase (Genschik *et al.*, 1997; 1998). At the same time, cloning experiments have revealed that the RNA 3′-terminal phosphate cyclase is conserved among Eucarya, Bacteria and Archaea, suggesting that the enzyme performs an important function in RNA metabolism.

This article reviews properties of the human and bacterial cyclases, their mechanism of action and substrate specificity. Possible biological functions of the enzymes are also discussed.

PURIFICATION AND CLONING OF THE HUMAN CYCLASE cDNA

The cyclase has been purified from HeLa cell extracts more than 6000-fold to approximately 80% homogeneity, with the final yield of 6% (Vicente & Filipowicz, 1988; Filipowicz & Vicente, 1990; Genschik *et al.*, 1997). The procedure involves DEAE-cellulose, Heparin-Sepharose, Poly(A)-Sepharose, Mono-S and Blue-Sepharose steps. Chromatography on Heparin-Sepharose yielded two peaks of activity (Filipowicz *et al.*, 1985); it is not known whether they represent two different cyclases known to be expressed in human cells (see below) or modified forms of the same protein. Covalent labelling of the enzyme with [α-^{32}P]ATP greatly facilitated identification of the cyclase polypeptide during final stages of purification. Analysis of the Blue-Sepharose fraction by the glycerol gradient centrifugation indicated that cyclase activity cosediments with the about 40-kDa band of adenylylated protein as determined by SDS/polyacrylamide gel electrophoresis followed by silver staining and autoradiography (Fig. 1). The purified cyclase did not contain detectable DNase, RNase, ATPase, GTPase, unspecific phosphatase or nucleoside diphosphokinase activities (Vicente & Filipowicz, 1988). In addition, Reinberg *et al.* (1985) have reported that partially purified cyclase from HeLa cells did not contain myokinase, 2′,3′-cyclic phosphodiesterase or RNA ligase activities. Likewise, no RNA ligase activity was associated with the purified recombinant human enzyme (Genschik *et al.*, 1997).

Purified protein was subjected to tryptic digestion and four peptides were microsequenced. Different combinations of degenerate oligonucleotide primers were used in order to clone, using a PCR approach, a cDNA encoding the enzyme. A partial PCR-derived cDNA was used as a probe to screen a HeLa λgt11 cDNA library. Eight positive clones were isolated; additional procedures were used to obtain the missing 5′-proximal cDNA sequence. The characterised cDNA encods a 39.4-kDa protein of 366 amino acids. All the microsequenced peptides are present in the deduced protein (Genschik *et al.*, 1997).

ACTIVITY OF THE HUMAN CYCLASE

The cyclase was overexpressed in *E. coli* as a fusion protein with glutathione *S*-transferase and purified using glutathione-Sepharose 4B resin. The recombinant protein was demonstrated to have RNA 3′-terminal phosphate cyclase activity. Since properties of the native and recombinant enzymes are very similar (Genschik *et al.*, 1997), they will be discussed together.

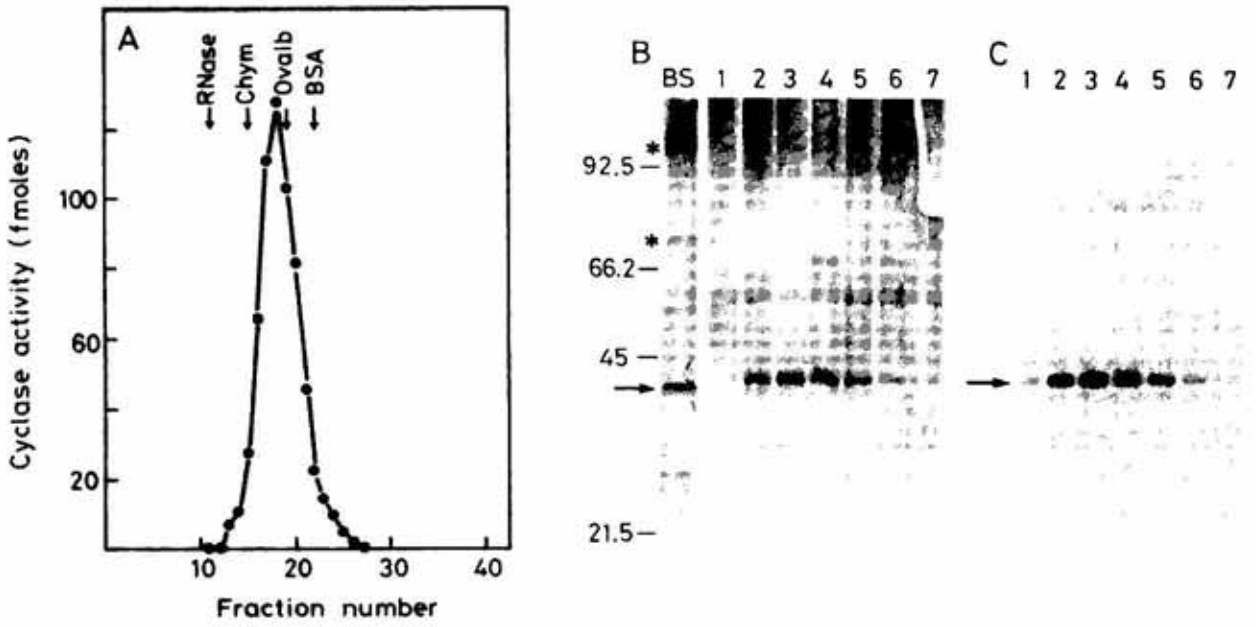

Figure 1. Analysis of the purified HeLa cell cyclase by glycerol gradient centrifugation.

(A) Profile of cyclase activity. One-half microgram of the Blue-Sepharose (BS-600)-purified fraction was analyzed by sedimentation in a 15–35% glycerol gradient. Centrifugation was for 36 h at 56000 r.p.m. in a Beckman SW60 rotor. Aliquots of 0.4 μl of each fraction were assayed for cyclase activity with 240 fmol of $(Np)_nGp^*$ substrate. The positions of marker proteins are indicated by arrows. (B and C) SDS/10% PAGE analysis of the gradient fractions after labelling with [α-^{32}P]ATP. Aliquots (60 μl) of glycerol gradient fractions 15 (lane 1) and 17–22 (lanes 2 to 7) were incubated with [α-^{32}P]ATP (40 μM, specific activity 2 Ci/mmol) for 3 h under standard conditions but without the oligonucleotide substrate; lane BS contained 20 ng of BS-600 fraction that was not incubated with [α-^{32}P]ATP. M_r ($\times 10^{-3}$) of marker proteins are shown at left. The band corresponding to cyclase is indicated by an arrow and the contaminating proteins in fraction BS-600 by asterisks. The artifactual bands visible in all gel lines (B) are caused by overstaining with silver. (B) Silver-stained gel. (C) Autoradiogram. (From Vicente & Filipowicz, 1988; reprinted with permission).

Cyclase assays

Activity of the human cyclase to catalyse conversion of the 3′-terminal monoester into cyclic phosphodiester has been measured either directly or by using a competition assay. Formation of the 2′,3′-cyclic phosphate in the RNA or oligoribonucleotide substrate can be directly assayed by thin-layer chromatography (TLC), following digestion of the reaction products with nuclease P1. Alternatively, since cyclisation renders the 3′-terminal phosphate resistant to the action of alkaline phosphatase, generation of the phosphatase-resistant radioactivity can be conveniently measured by retention on activated charcoal (Norit). Either assay requires substrates in which the 3′-terminal phosphate is the only labelled phosphate group, such as oligoribonucleotides CCCCACCCCG$^{3'}$p* or AAAAUAAAAG$^{3'}$p* (p* denotes ^{32}P-labelled phosphate), which has been used in most of the experiments. For the TLC assay, substrates labelled at the penultimate phosphate such as CCCCACCCCGp*C$^{3'}$p or AAAAUAAAAGp*C$^{3'}$p, conveniently prepared by ligation of the [5′-^{32}P]pCp to the 3′-hydroxyl-terminated RNA, can also be used. In competition assays, different unlabelled nucleotides or oligonucleotides are tested for their ability to compete with the cyclisation of the radiolabelled AAAAUAAAAG$^{3'}$p* substrate. Finally, another assay measures activity of different compounds to release AMP from the preformed adenylylated enzyme complex. The complex is formed by preincubation of cyclase

with [α-^{32}P]ATP and incubations are then continued in the presence of increasing quantities of different oligonucleotides.

Substrate specificity

The human cyclase can utilize molecules with different sequences and base composition as substrates. A nature of the 3′-terminal base also does not appear to be of importance. Molecules demonstrated to act as substrates include: tRNA-pNp, 5S rRNA-pCp, tobacco mosaic virus RNA fragments modified by ligation of pGp, pAp or pCp, and pure oligoribonucleotides of different lenghts and sequences such as CCCCACCCCGp, CCCCCACCCGCp, AAAAUAAAAGp, AAAAUAAAAGCp, $(Up)_{10}$pGp, $(Ap)_n$pAp or AUGp. Ribonucleoside 3′-phosphates and ribonucleoside 5′,3′-diphosphates do not act as substrates for the cyclase (Filipowicz *et al.*, 1983; 1985; Reinberg *et al.*, 1985; Filipowicz & Vicente, 1990; Genschik *et al.*, 1997). Since 3′-phosphorylated dinucleotides do not compete with the phosphate cyclisation in AAAAUAAA - $AG^{3'}p^{*}$ (Genschik *et al.*, 1997), it appears that trinucleotides are the shortest oligonucleotides able to act as substrates for the cyclase.

The observation that oligoribonucleotides containing terminal 2′-deoxy- or 2′-*O*-methylribose can be converted, in the presence of cyclase and ATP, into products bearing the 3′-terminal structures $dN^{3'}pp^{5'}A$ and $N^{m3'}pp^{5'}A$, respectively (Filipowicz *et al.*, 1985), raised the possibility that 3′-phosphorylated DNA molecules can be physiological substrates for the cyclase. However, comparison of the ability of the 3′-phosphorylated oligoribonucleotide, $AAAAUAAAAG^{3'}p$, and the oligodeoxyribonucleotide of equivalent sequence, $AAAATAAAAG^{3'}p$, to act as substrates in the cyclisation assay indicated that the DNA version is approximately 500-fold less active (Fig. 2). Likewise, a mixture of 3′-phosphorylated oligodeoxyribonucleotides [$(dN)_npdN^{3'}p$] was a poor competitor in the reaction. The 3′-phosphorylated oligoribo- and oligodeoxyribonucleotides were also compared for their ability to release AMP from the preformed adenylylated enzyme complex. Again, deoxyribooligonucleotides were found to be at least a 1000-fold less active than the RNA compounds (Fig. 2). Taken together, these results indicate that 3′-phosphorylated oligodeoxyribonucleotides are very inefficient substrates for the cyclase.

Requirements of the cyclization reaction

The cyclase requires ATP as a co-factor but it can also use, though very inefficiently, other ribonucleoside triphosphates. The apparent K_m values for ATP and GTP are 6 μM and 200 μM, respectively. ADP and nonhydrolyzable α,β-methylene (AMPCPP) and β,γ-methylene (AMPPCP) analogs of ATP, do not act as co-factors. Adenosine 5′-(γ-thio)triphosphate (ATPγS) is a slightly more effective co-factor than ATP. The optimal concentration of monovalent ions, Na^+ or K^+, is 150–200 mM and of divalent ions, Mg^{2+}, is 4–5 mM. Only about 5% of activity is seen when Mg^{2+} is replaced with Mn^{2+} or Ca^{2+}. The pH optimum for the human cyclase is 8.0–9.0 (Filipowicz *et al.*, 1985; Reinberg *et al.*, 1985; Vicente & Filipowicz, 1988; Genschik *et al.*, 1997).

INTRACELLULAR LOCALIZATION OF THE CYCLASE AND ITS EXPRESSION IN DIFFERENT HUMAN TISSUES AND CELL LINES

Since specific antibodies against the human cyclase are not available, the intracellular localization of the enzyme was determined by an epitope-tagging and indirect immunofluorescence approach. A *myc* epitope was fused in frame to the N-terminus of the protein. Expression of the tagged protein was studied in transfected HeLa cells, rat glioma C6 cells and mouse embryonal carcinoma P19 cells. Using a mouse anti-*myc* monoclonal antibody and fluorescein isothiocyanate-conjugated

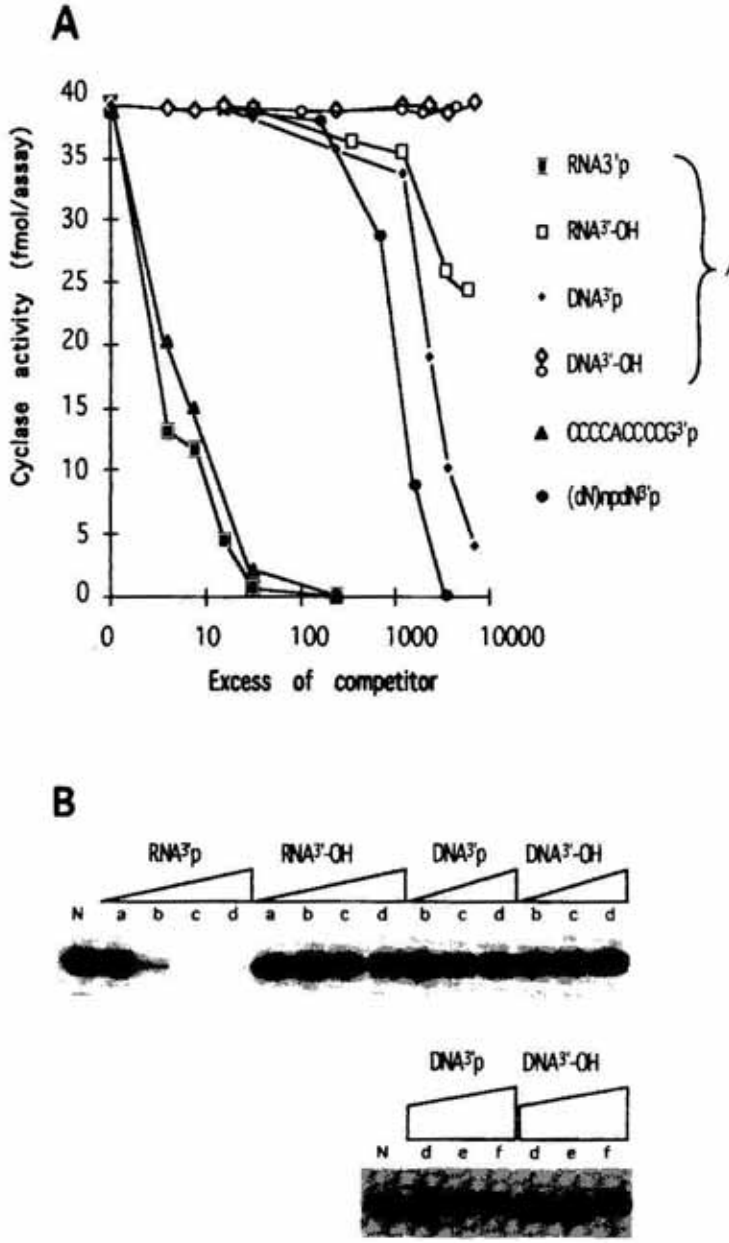

Figure 2. Activity of oligoribo- or oligodeoxyribonucleotides containing either 3′-p or 3′-OH termini in the competition (A) and AMP release (B) assays.

(A) Cyclase assays were performed as described by Genschik *et al.* (1997). They contained 360 pg of the overexpressed cyclase, 40 fmol of AAAAUAAAAG$^{3'}$p* and different unlabelled oligonucleotides, acting as substrates or competitors, at molar excess over AAAAUAAAAG$^{3'}$p* as indicated. The 3′-OH-containing AAAATAAAAG (DNA$^{3'}$OH) was obtained either by chemical synthesis (o) or by dephosphorylation of DNA$^{3'}$p with CIP (◇). (B) AMP release assays. Different amounts of substrates (described in panel A) were added: none (lane N); 2 fmol (lane a); 22 fmol (lane b); 220 fmol (lane c); 2200 fmol (lane d); 20 pmol (lane e) and 200 pmol (lane f). (Reproduced, with permission, from Genschik *et al.*, 1997).

(FITC-conjugated) goat anti-mouse antibody, and the confocal laser scanning microscopy, it was found that 98–99% of cyclase localizes to the nucleus. The protein shows a diffuse distribution throughout the nucleoplasm and is excluded from the nucleoli. The staining was independent of the amount of plasmid used for transfection or the time at which cells were collected after transfection (Genschik *et al.*, 1997).

Expression of the cyclase gene was studied by Northern analysis. The cyclase mRNA was found to be expressed in a wide range of human tissues (heart, brain, placenta, lung, skeletal muscle, kidney, pancreas) and cells lines (HeLa, MCF7 adenocarcinoma, K562 myeloid leukemia, Namalwa Burkitt lymphoma, HUT78 and Jurkat T-cell lymphomas, and 293T primary embryonic kidney line). Among the tissues analyzed, the highest mRNA level was observed in a skeletal muscle. Two hybridizing RNA species, approximately 1.8-kb- and 3-kb-long, were identified in all tissues and cell lines. The 1.8-kb mRNA corresponds in size to the characterized cDNA; identity of the longer RNA remains to be established (Genschik *et al.*, 1997).

PROPERTIES OF THE *E. COLI* CYCLASE

Cloning of the human cDNA has revealed that genes encoding proteins of unknown function, but sharing significant similarity with the human cyclase, are present in many different organisms belonging to all three domains of life: Eucarya, Bacteria and Archaea (see below). The protein encoded by the *E. coli* gene, having similarity with the human cyclase cDNA, has been bacterially overexpressed as a fusion protein with the histidine tag added to the C-terminus for the purification purpose (Fig. 3). Analysis of the requirements and substrate specificity of the purified recombinant *E. coli* protein demonstrated that properties of the bacterial and human enzymes are very similar (Genschik *et al.*, 1997; 1998).

The *E. coli* enzyme can catalyze cyclization of the 3′-terminal phosphate in synthetic oligoribonucleotides such as CCCCACCCCGp, AAAAUAAAAGp (Fig. 3) and AAAAUAAAAGCp and in natural RNAs (the *E. coli* 5S rRNA or human U14 snoRNA) modified by ligation of [5′-^{32}P]pCp to the 3′-terminus. Competition experiments carried out with nucleoside 3′-monophosphates and nucleoside 5′,3′-bisphosphates indicated that these compounds do not act as substrates for the *E. coli* enzyme. Furthermore, competition and AMP release experiments have shown that 3′-phosphate-terminated DNA molecules are two to three orders of magnitude poorer substrates than RNAs.

Although requirements of the *E. coli* cyclase are similar to that of the human protein, some differences are also apparent. One of them is the ability of Mn^{2+} ions to replace Mg^{2+} in reactions catalyzed by the *E. coli* but not by the human cyclase. In the presence of Mn^{2+}, the enzyme activity was 50–70% higher than in the presence of Mg^{2+}. With both cations, a broad optimum was found at 1–4 mM. No activity was seen with Ca^{2+}, Zn^{2+}, or Cu^{2+}.

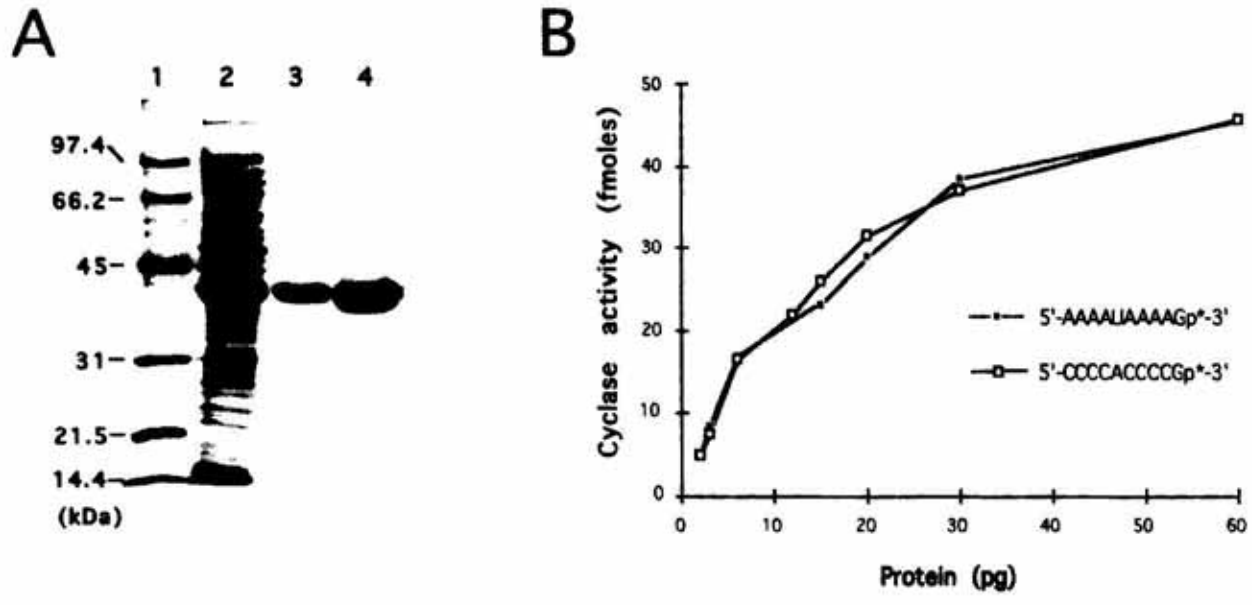

Figure 3. Purification and activity of the overexpressed *E. coli* cyclase.

(A) Purification of the protein fused to the 6 × His tag as monitored by SDS/PAGE. Lane 1, size markers; lane 2, 40 μg of extract from the induced *E. coli* culture; lane 3 and 4, 3 μg and 9 μg of the purified protein, respectively. (B) Comparison of CCCCACCCCG$^{3'}$ p* and AAAAUAAAAG$^{3'}$ p* substrates. (Reproduced, with permission, from Genschik *et al.*, 1997).

Activity of the *E. coli* enzyme was similar at 0 and 0.1 M NaCl, and addition of NaCl to 0.2 M or 0.4 M inhibited cyclisation by 30% and 70%, respectively. At 20 or 50 mM, Na-phosphate did not inhibit the reaction but Na-pyrophosphate was strongly inhibitory (Genschik *et al.*, 1998). The *E. coli* enzyme preferentially uses ATP as a cofactor. However, GTP is a relatively more efficient cofactor with the *E. coli* than with the human enzyme. Apparent K_m values for ATP and GTP are, respectively, 20 and 100 μM for the *E. coli* protein and 6 μM and 200 μM for the human protein. ATP could not be replaced by either AMPCPP or AMPPCP, or by the β,γ-imido analog of ATP (AMPPNP). ATPγS was about 20% more active than ATP (Genschik *et al.*, 1998).

Consistent with the cyclization reaction proceeding *via* the formation of the covalent enzyme-NMP intermediate, the bacterial enzyme was found to undergo adenylylation and the adenylyl group could be released from the preformed cyclase-AMP complex protein upon incubation with the 3′-phosphorylated RNA but not the 3′-OH-terminated RNA. Covalent labelling of the bacterial cyclase with [α-^{32}P]GTP and, much less efficiently, with [α-^{32}P]CTP, [α-^{32}P]UTP and [α-^{32}P]dATP was also observed (Genschik *et al.*, 1998).

STRUCTURE OF THE CYCLASE OPERON IN *E. COLI*

The *E. coli* cyclase gene, named *rtcA*, forms part of a previously uncharacterized operon, positioned at 76 min on the *E. coli* K12 chromosome and containing two additional open reading frames (ORFs) (Fig. 4). The ORF positioned immediately upstream of the cyclase, named *rtcB*, encodes a protein which is also highly conserved between Eucarya, Bacteria, and Archaea. Analysis of sequences of RtcB and RtcB-like proteins from other organisms, did not provide any clue as to the function of the protein. Another ORF, called *rtcR*, is positioned upstream of the *rtcA/rtcB* transcription unit and is transcribed in the opposite direction. It encodes a protein having features of regulators involved in a transcriptional control of operons depending on the alternative sigma factor, σ^{54}. By overexpressing the N-terminally truncated form of RtcR, it has been demonstrated that this regulator indeed controls expression of the cyclase (*rtcA*) and *rtcB* genes in a σ^{54}-dependent manner. Physiologically, activation of most σ^{54}-spe-cific regulators, and consequently of operons controlled by them, is brought about either by phosphorylation or by binding of specific effector molecules to the regulator (reviewed by Shingler, 1996). A potential effector interacting with RtcR is not known. Its identification might provide some clues as to the function of the cyclase and its operon in *E. coli*. With the exception of *Myxococcus xantus* (Keseler & Kaiser, 1997), in all bacteria studied to date, σ^{54} and consequently also genes or operons dependent on it were found not to be essential for growth. Consistent with this, knock-out experiments have shown that the gene encoding cyclase is not required for *E. coli* growth on either LB or a minimal M9 medium (Genschik *et al.*, 1998). Products of genes controlled by σ^{54} factors have very diverse physiological functions. They are involved in specialized metabolic processes such as utilization of alternative carbon sources or assimilation and fixation of nitrogen (reviewed by Merrick, 1993; Shingler, 1996). In *Caulobacter crescentus* and *Myxococcus xantus*, σ^{54} and σ^{54}-specific transcription regulators are involved in important cell differentiation decisions (Gulati *et al.*, 1995; Wingrove & Gober, 1994; Keseler & Kaiser, 1995, and references therein). Identification of physiological conditions leading to the activation of the cyclase operon would greatly help to establish a role of this enzyme in RNA metabolism in bacteria.

CONSERVATION OF CYCLASE AMONG EUCARYA, BACTERIA AND ARCHAEA

To date, genes or Expressed Sequence Tags (ESTs) encoding cyclase-like proteins have been identified in organisms belonging to all three kingdoms. Eukaryotic organisms include mouse, *Drosophila melanogaster*, zebra fish, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*, *Toxoplasma gondii* and *Dictyostelium discoideum*. In mammals, at least two genes expressing related proteins, showing approximately 30% identity and 52% similarity, have been identified (Genschik *et al.*, 1997). In bacteria, genes encoding cyclase-like proteins, in addition to *E. coli*, are present in *Pseudomonas aeruginosa* and *Aquitex aeolicus*. They are also found in *Methanococcus jannaschii* and other Archaea. However, no obvious homologs to the cyclase could be found in the streamlined bacterial genomes such as that of *Haemophilus influenzae*, *Mycoplasma genitalium* and *Helicobacter pylori*, and also in *Bacillus subtilis* and *Synechocystis* sp. (Genschik *et al.*, 1997; unpublished observations).

Proteins belonging to the family of RNA cyclases have no apparent structural features in common with other known proteins. In particular, no sequences corresponding to known

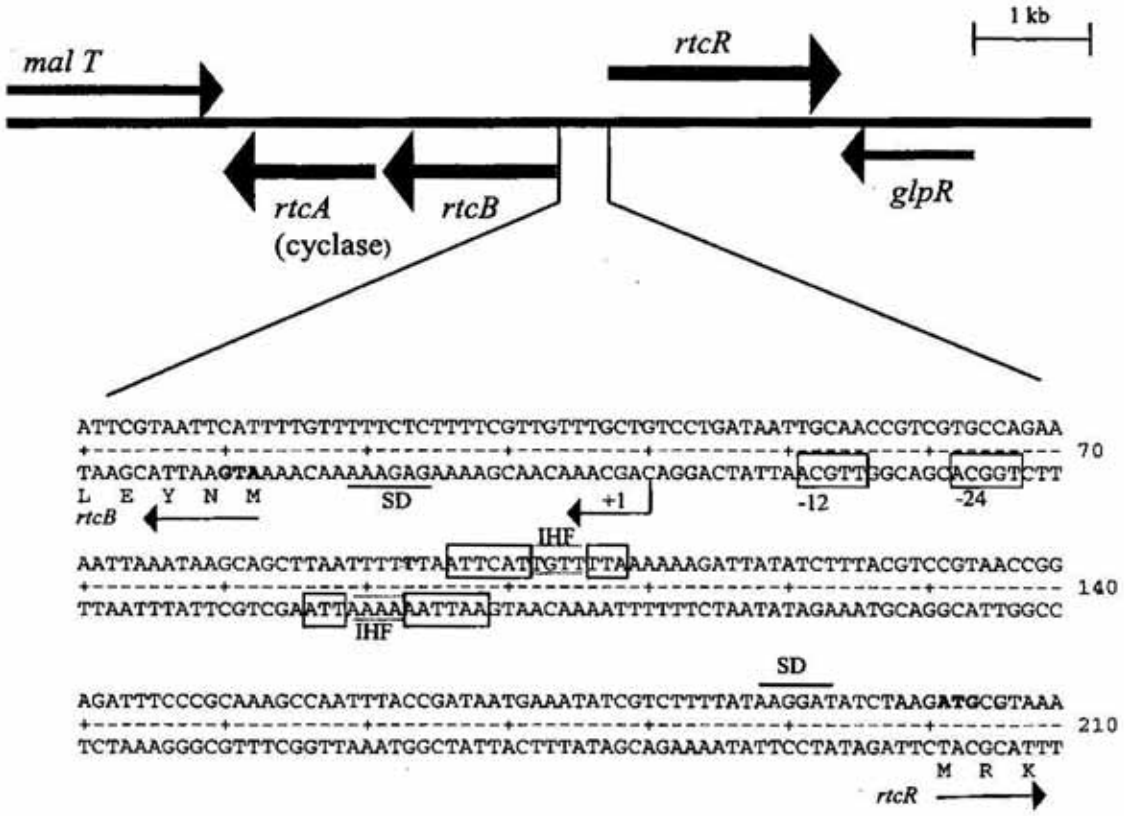

Figure 4. Structure of the cyclase operon and the sequence of the intergenic promoter region between *rtcB* and *rtcR*.

Genes constituting the cyclase operon are shown as filled black arrows. The *rtcB* transcription start site, as established by primer extension analysis, is indicated by a bent arrow. Translation start sites for *rtcB* and *rtcR* are marked with arrows and putative Shine-Dalgarno sequences are underlined. The −12 and −24 consensus sequences characteristic of σ^{54} promoters are boxed. Sequences resembling the most highly conserved regions of the consensus binding site, $AATCAAN_4TTA$, for the integration host factor (IHF) are indicated by boxes connected by the dotted lines. (Reproduced, with permission, from Genschik *et al.*, 1998).

nucleotide-binding motifs or RNA-binding domains could be identified. Previous analysis of the stability of the covalent cyclase-AMP complex suggested that AMP is linked to the protein *via* a phosphoamide linkage, possibly involving the ε-amino group of a lysine (Reinberg *et al.*, 1985; Vicente & Filipowicz, 1988). However, no lysine is present at conserved position in all proteins. With respect to the ability to transfer the nucleotidyl group from the protein-NMP intermediate to the terminal phosphate or pyrophosphate in nucleic acids, the cyclase resembles RNA and DNA ligases and capping enzymes (reviewed by Shuman & Schwer, 1995), but none of the sequence motifs shared by these enzymes (Shuman & Schwer, 1995) is apparent in the RNA cyclase.

POSSIBLE FUNCTION OF THE CYCLASE

The biological role of the RNA 3′-terminal phosphate cyclase remains unknown. However, conservation of the enzyme among Eucarya, Bacteria and Archaea strongly argues that it performs an important function in RNA metabolism. The cyclase likely functions in some aspects of cellular RNA processing. The predominantly nucleoplasmic localization of the enzyme in mammalian cells and its ability to use poly- and oligoribonucleotides but not mono- or diribonucleotides as substrates are consistant with this possibility. Experiments in which activity of oligoribo- and oligodeoxyribonucleotides of identical sequence was compared have revealed that oligoribonucleotides are about 500-fold better substrates than oligodeoxyribonucleotides for the human and the *E. coli* cyclase. Hence, it is rather unlikely that the cyclase participates in DNA- rather than RNA-related processes.

The anabolic function of the 2′,3′-cyclic phosphate in RNA first emerged when it was found that eukaryotic RNA ligases require 2′,3′-cyclic ends for RNA ligation (Konarska *et al.*, 1981; 1982; Filipowicz *et al.*, 1983; Filipowicz & Shatkin, 1983; Furneaux *et al.*, 1983; Greer *et al.*, 1983a; reviewed by Filipowicz & Gross, 1984; Westaway & Abelson, 1995). This requirement applies to both of the nonorganellar RNA ligases characterized to date, one ligating RNA ends *via* the unusual 3′,5′-phosphodiester, 2′-phosphomonoester linkage, and the other joining the ends *via* the regular 3′,5′-phosphodiester. The involvement of the two RNA ligases in nuclear pre-tRNA splicing is well documented but these enzymes might also function in the ligation of other natural RNA molecules such as virusoids and viroids (reviewed by Filipowicz & Gross, 1984; Phizicky & Greer, 1993; Westaway & Abelson, 1995; see also Sidrauski & Walter, 1997). Although the splicing endonucleases directly generate 5′-tRNA halves carrying 2′,3′-cyclic phosphate during tRNA splicing reactions, it is possible that other putative substrates depend upon the action of RNA 3′-terminal phosphate cyclase to form the cyclic ends. It is of interest that the only cellular RNA ligase identified to date in bacteria also requires 2′,3′-cyclic ends for ligation (Greer *et al.*, 1983b; Arn & Abelson, 1996). The bacterial ligase joins RNA ends *via* the 2′,5′-phosphodiester linkage. The *E. coli* enzyme appears to be specific for ligation of tRNA halves (Arn & Abelson, 1996) but its physiological substrates and the mechanism of production of their cyclic ends remain unknown. tRNA genes in Archaea contain introns which structurally resemble nuclear tRNA gene introns but the requirements for ligation of tRNA halves in Archaea have not been characterized (reviewed by Westaway & Abelson, 1995).

Alternatively, the cyclase might be responsible for cyclization of the 3′-terminal phosphate in the spliceosomal U6 snRNA. Lund & Dahlberg (1992) have found that U6 RNA in most eukaryotes investigated contains a 2′,3′-cyclic phosphate end. The mechanism and enzymes responsible for this modification are unknown. Finally, it is possible that the cy-

clase is not a component of an RNA ligation pathway but modifies the RNA 3′ end for a different purpose, for example, activation of RNA for exonucleolytic degradation. Richards & Laskowski (1969) have shown that the presence of a terminal 3′-phosphate makes an oligonucleotide largely resistant to 3′→ 5′ exonucleolytic degradation by snake venom phosphodiesterase. The resistance may be due to the strong negative charge at the 3′ end since oligonucleotides bearing a 2′,3′-cyclic phosphate were found to be almost as active as the 3′-OH-terminated substrates. Identity of physiological substrates for the RNA 3′-terminal phosphate cyclase and its biological role remain to be established.

REFERENCES

Arn, E.A. & Abelson, J.N. (1996) The 2′-5′ RNA ligase of *Escherichia coli*. Purification, cloning and genomic disruption. *J. Biol. Chem.* **269**, 31145–31153.

Filipowicz, W. & Shatkin, A.J. (1983) Origin of splice junction phosphate in tRNAs processed by HeLa cell extract. *Cell* **32**, 547–557.

Filipowicz, W. & Gross, H.J. (1984) RNA ligation in eukaryotes. *Trends Biochem. Sci.* **9**, 68–71.

Filipowicz, W., Konarska, M., Gross, H.J. & Shatkin, A.J. (1983) RNA 3′-terminal phosphate cyclase activity and RNA ligation in HeLa cell extract. *Nucleic Acids Res.* **11**, 1405–1418.

Filipowicz, W., Strugala, K., Konarska, M. & Shatkin, A.J. (1985) Cyclization of RNA 3′-terminal phosphate by cyclase from HeLa cells proceeds *via* formation of $N(^{3'})pp(^{5'})A$ activated intermediate. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1316–1320.

Filipowicz, W. & Vicente, O. (1990) RNA 3′-terminal phosphate cyclase from HeLa cells. *Methods Enzymol.* **181**, 499–510.

Furneaux, H., Pick, L. & Hurwitz, J. (1983) Isolation and characterization of RNA ligase from wheat germ. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3933–3937.

Genschik, P., Drabikowski, K. & Filipowicz, W. (1998) Characterization of the *Escherichia coli* RNA 3′-terminal phosphate cyclase and its σ^{54}-regulated operon. *J. Biol. Chem.* **273**, 25516–25526.

Genschik, P., Billy, E., Swianiewicz, M. & Filipowicz, W. (1997) The human RNA 3′-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea. *EMBO J.* **16**, 2955–2967.

Greer, C.L., Peebles, C.L., Gegenheimer, P. & Abelson, J. (1983a) Mechanism of action of a yeast RNA ligase in tRNA splicing. *Cell* **32**, 537–546.

Greer, C.L., Javor, B. & Abelson, J. (1983b) RNA ligase in Bacteria: Formation of a 2′,5′ linkage by an *E. coli* extract. *Cell* **33**, 899–906.

Gulati, P., Xu, D. & Kaplan, H.B. (1995) Identification of the minimum regulatory region of a *Myxococcus xanthus* A-signal-dependent developmental gene. *J. Bacteriol.* **177**, 4645–4651.

Keseler, I.M. & Kaiser, D. (1995) An early A-signal-dependent gene in *Myxococcus xanthus* has a σ^{54}-like promoter. *J. Bacteriol.* **177**, 4638–4644.

Keseler, I.M. & Kaiser, D. (1997) σ^{54}, a vital protein for *Myxococcus xanthus*. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1979–1984.

Konarska, M., Filipowicz, W., Domdey, H. & Gross, H.J. (1981) Formation of a 2′-phosphomonoester, 3′,5′-phosphodiester linkage by a novel RNA ligase in wheat germ. *Nature* **293**, 112–116.

Konarska, M., Filipowicz, W. & Gross, H.J. (1982) RNA ligation *via* 2′-phosphomonoester, 3′-, 5′-phosphodiester linkage: Requirement of 2′, 3′-cyclic phosphate termini and involvement of a 5′-hydroxyl polynucleotide kinase. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1474–1478.

Lund, E. & Dahlberg, J.E. (1992) Cyclic 2′,3′-phosphates and nontemplated nucleotides at the 3′ end of spliceosomal U6 small nuclear RNA's. *Science* **255**, 327–330.

Merrick, M.J. (1993) In a class of its own – the RNA polymerase sigma factor σ^{54} (σ^{N}). *Mol. Microbiol.* **10**, 903–909.

Phizicky, E.M. & Greer, C. (1993) Pre-tRNA splicing: Variation on a theme or exception to the rule? *Trends Biochem. Sci.* **18**, 31–34.

Reinberg, D., Arenas, J. & Hurwitz, J. (1985) The enzymatic conversion of 3′-phosphate terminated RNA chains to 2′,3′-cyclic phosphate derivatives. *J. Biol. Chem.* **260**, 6088–6097.

Richards, G.M. & Laskowski, M., Sr. (1969) Negative charge at the 3′ terminus of oligonucleotides and resistance to venom exonuclease. *Biochemistry* **8**, 1786–1795.

Schwartz, R.C., Greer, C.L., Gegenheimer, P. & Abelson, J. (1983) Enzymatic mechanism of an RNA ligase from wheat germ. *J. Biol. Chem.* **258**, 8374–8383.

Sekiguchi, J. & Shuman, S. (1997) Site-specific ribonuclease activity of eukaryotic topoisomerase I. *Mol. Cell* **1**, 89–97.

Shingler, V. (1996) Signal sensing by σ^{54}-dependent regulators: Derepression as a control mechanism. *Mol. Microbiol.* **19**, 409–416.

Shuman, S. & Schwer, B. (1995) RNA capping enzyme and DNA ligase: A superfamily of covalent nucleotidyl transferases. *Mol. Microbiol.* **17**, 405–410.

Sidrauski, C. & Walter, P. (1997) The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90**, 1031–1039.

Symons, R.H. (1992) Small catalytic RNAs. *Annu. Rev. Biochem.* **61**, 641–671.

Vicente, O. & Filipowicz, W. (1988) Purification of RNA 3′-terminal phosphate cyclase from HeLa cells. *Eur. J. Biochem.* **176**, 431–439.

Westaway, S.K. & Abelson, J. (1995) Splicing of tRNA precursors; in *tRNA: Structure, Biosynthesis and Function* (Soell, D. & RajBhandary, U., eds) pp. 79–92, American Society for Microbiology, Washington, D.C.

Wingrove, J.A. & Gober, J.W. (1994) A σ^{54} transcriptional activator also functions as a pole-specific repressor in *Caulobacter*. *Genes Dev.* **8**, 1839–1852.