The 2′-5′ RNA Ligase of Escherichia coli

PURIFICATION, CLONING, AND GENOMIC DISRUPTION

(Received for publication, August 7, 1996, and in revised form, September 18, 1996)

Eric A. Arn‡ and John N. Abelson

From the Division of Biology 147-75, California Institute of Technology, Pasadena, California 91125

An RNA ligase previously detected in extracts of Escherichia coli is capable of joining Saccharomyces cerevisiae tRNA splicing intermediates in the absence of ATP to form a 2′-5′ phosphodiester linkage (Greer, C., Javor, B., and Abelson, J. (1983) Cell 33, 899–906). This enzyme specifically ligates tRNA half-molecules containing nucleoside base modifications and shows a preference among different tRNA species. In order to investigate the function of this enzyme in RNA metabolism, the ligase was purified to homogeneity from E. coli lysate utilizing chromatographic techniques and separation of proteins by SDS-polyacrylamide gel electrophoresis. A single polypeptide of approximately 20 kdaltons exhibited RNA ligase activity. The amino terminus of this protein was sequenced, and the open reading frame (ORF) encoding it was identified by a data base search. This ORF, which encodes a novel protein with a pre-

editing in kinetoplastids (5). At least two other species of RNA ligase, the metazoan-specific “animal pathway” RNA ligase (6) and the archaeal stable RNA splicing ligase (7), have been identified, but the precise in vivo substrate(s) and function(s) of these enzymes have yet to be determined.

The existence of RNA ligase in bacteria in the absence of bacteriophage infection was discovered in this laboratory in 1983 (8). An activity capable of performing the ligation step of eukaryotic tRNA splicing was detected in extracts of a wide range of bacteria including members of the Alpha and Gamma subdivisions of proteobacteria, green sulfur bacteria, and low G + C content Gram-positive bacteria (8). In extracts of E. coli the ligase activity had a substrate specificity restricted to 4 of the 10 Saccharomyces cerevisiae tRNA splicing intermediates: tRNA Tyr, Phe, Lys2, and Trp half-molecules. The reaction mechanism of the E. coli RNA ligase apparently differed from that of known RNA ligases since it did not require a nucleoside triphosphate cofactor, and the product contained an unusual 2′-5′ phosphodiester bond at the ligated junction (8).

The discovery of RNA ligase in E. coli implies the existence of a novel form of bacterial RNA processing. No intervening sequences of the type found in eukaryotic nuclear or archaeal tRNA genes (which require enzymatic excision and religation) occur in known bacterial tRNA genes, although self-splicing introns are found in the tRNA genes of certain cyanobacteria (9) and some proteobacteria (10). In fact, no introns of any kind are found in the full genomic complement of tRNA genes in E. coli, Mycoplasma capricolum, and Hemophilus influenzae (11–13). To date no RNA processing event that would require the action of an RNA ligase enzyme has been observed to occur in any bacteria, indicating that elucidation of the substrate and function of the 2′-5′ RNA ligase should reveal a previously unknown step in bacterial RNA metabolism.

A genuine in vivo function for the E. coli RNA ligase activity observed in vitro is suggested by the occurrence of 2′-5′ linkages in native E. coli RNA. Several forms of 2′-5′-linked oligoadenylates have been detected in acid-soluble extracts of E. coli (14). The most abundant species of these oligoribonucleotides observed was 2′-5′-linked adenosine dinucleotide 3′-monophosphate, which was estimated to exist at an intracellular concentration over 100 nm. This dinucleotide could not be an intermediate in oligoadenylate synthesis, and it has been suggested that this species may be a degradation product of RNAs containing individual 2′-5′ bonds among standard 3′-5′ linkages (14). Since the formation of 2′-5′ linkages is not catalyzed by RNA polymerase, these bonds must be added in a posttranscriptional processing event by an enzyme such as the 2′-5′ RNA ligase.

In order to study the function and mechanism of the 2′-5′ RNA ligase, the enzyme was purified to homogeneity from E. coli extracts. A single polypeptide was found to contain RNA ligase activity. This protein was partially sequenced, and the
ORF encoding it was identified. This ORF was amplified from E. coli genomic DNA and cloned. The chromosomal loci containing the ligase gene was disrupted, abolishing ligase activity in cellular extracts. Cells completely lacking ligase activity grow similarly to the parent strain. E. coli strains overexpressing recombinant 2'-5' RNA ligase are temperature-sensitive for growth. The ligase reaction was also studied using purified enzyme and was found to be reversible in vitro.

**EXPERIMENTAL PROCEDURES**

*Strains—*E. coli strains utilized were HB101, XL1-Blue (Novagen), and HSD947, DH110B, RecA-.

**Preparation of RNA Substrates for Ligation—**RNA polymerase III transcription of pre-tRNA<sup>79</sup> in yeast extract was performed by the method of Evans and Engelke (15), using extracts prepared as described. The template for Pol III transcription of pre-tRNA<sup>79</sup> was supercoiled pYSUP6 (16). T7 RNA polymerase transcription was performed as described by Sampson and Saks (17). pre-tRNA<sup>79</sup> was transcribed from the pre-tRNA<sup>79</sup> gene of the pUC19 plasmid. The T7 promoter was added at the 3' end of the pre-tRNA<sup>79</sup> gene as alter the acceptor stem base pairs. C<sup>3</sup>G → G<sup>7</sup>T, G<sup>7</sup>T→ G<sup>7</sup>A, A<sup>7</sup>G→ C, and G<sup>8</sup>→ C to improve transcription yields. This was cloned into the plBetaScript vector (Strategene). tRNA substrates were modified by incubation at 24°C for 45 min in the same extract and reaction conditions as were used for Pol III transcription. tRNA precursors were cleaved as described by Peebles et al. (2), using partially purified S. cerevisiae tRNA splicing endonuclease fractions from the hydroxypapatite step or later as described in the Raabt protocol (2, 19).

All RNA transcripts were gel purified by polyacrylamide gel electrophoresis (PAGE) in 1 x TBE (Tris borate EDTA buffer), 7 M urea and visualized by autoradiography. RNAs were eluted from crushed gel slices in 0.6 M NH<sub>4</sub>OAc, 2 mM EDTA, 0.005% Nonidet P-40 at room temperature with vortexing for 20 min. RNA eluates were extracted with phenol/chloroform (1:1, pH 4.5) and chloroform and then ethanol precipitated in the presence of glycogen and resuspended in distilled water.

**RNA Ligation Assay—**RNA half-molecules were annealed for ligation in 2 x ligation buffer (1 x 40 mM HEPES, pH 7.8, 3 mM MgCl<sub>2</sub>, 2 mM spermidine, 5% glycerol) by heating to 85°C and slowly cooling to 70°C. The ligase reaction was also studied using purified enzyme and was found to be reversible in vitro.

**Preparation of RNA Substrates for Ligation—**RNA polymerase III transcription of pre-tRNA<sup>79</sup> in yeast extract was performed by the method of Evans and Engelke (15), using extracts prepared as described. The template for Pol III transcription of pre-tRNA<sup>79</sup> was supercoiled pYSUP6 (16). T7 RNA polymerase transcription was performed as described by Sampson and Saks (17). pre-tRNA<sup>79</sup> was transcribed from the pre-tRNA<sup>79</sup> gene of the pUC19 plasmid. The T7 promoter was added at the 3' end of the pre-tRNA<sup>79</sup> gene as alter the acceptor stem base pairs. C<sup>3</sup>G → G<sup>7</sup>T, G<sup>7</sup>T→ G<sup>7</sup>A, A<sup>7</sup>G→ C, and G<sup>8</sup>→ C to improve transcription yields. This was cloned into the plBetaScript vector (Strategene). tRNA substrates were modified by incubation at 24°C for 45 min in the same extract and reaction conditions as were used for Pol III transcription. tRNA precursors were cleaved as described by Peebles et al. (2), using partially purified S. cerevisiae tRNA splicing endonuclease fractions from the hydroxypapatite step or later as described in the Raabt protocol (2, 19).

All RNA transcripts were gel purified by polyacrylamide gel electrophoresis (PAGE) in 1 x TBE (Tris borate EDTA buffer), 7 M urea and visualized by autoradiography. RNAs were eluted from crushed gel slices in 0.6 M NH<sub>4</sub>OAc, 2 mM EDTA, 0.005% Nonidet P-40 at room temperature with vortexing for 20 min. RNA eluates were extracted with phenol/chloroform (1:1, pH 4.5) and chloroform and then ethanol precipitated in the presence of glycogen and resuspended in distilled water.

**RNA Ligation Assay—**RNA half-molecules were annealed for ligation in 2 x ligation buffer (1 x 40 mM HEPES, pH 7.8, 3 mM MgCl<sub>2</sub>, 2 mM spermidine, 5% glycerol) by heating to 85°C and slowly cooling to 70°C. The ligase reaction was also studied using purified enzyme and was found to be reversible in vitro.

**Preparation of RNA Substrates for Ligation—**RNA polymerase III transcription of pre-tRNA<sup>79</sup> in yeast extract was performed by the method of Evans and Engelke (15), using extracts prepared as described. The template for Pol III transcription of pre-tRNA<sup>79</sup> was supercoiled pYSUP6 (16). T7 RNA polymerase transcription was performed as described by Sampson and Saks (17). pre-tRNA<sup>79</sup> was transcribed from the pre-tRNA<sup>79</sup> gene of the pUC19 plasmid. The T7 promoter was added at the 3' end of the pre-tRNA<sup>79</sup> gene as alter the acceptor stem base pairs. C<sup>3</sup>G → G<sup>7</sup>T, G<sup>7</sup>T→ G<sup>7</sup>A, A<sup>7</sup>G→ C, and G<sup>8</sup>→ C to improve transcription yields. This was cloned into the plBetaScript vector (Strategene). tRNA substrates were modified by incubation at 24°C for 45 min in the same extract and reaction conditions as were used for Pol III transcription. tRNA precursors were cleaved as described by Peebles et al. (2), using partially purified S. cerevisiae tRNA splicing endonuclease fractions from the hydroxypapatite step or later as described in the Raabt protocol (2, 19).

All RNA transcripts were gel purified by polyacrylamide gel electrophoresis (PAGE) in 1 x TBE (Tris borate EDTA buffer), 7 M urea and visualized by autoradiography. RNAs were eluted from crushed gel slices in 0.6 M NH<sub>4</sub>OAc, 2 mM EDTA, 0.005% Nonidet P-40 at room temperature with vortexing for 20 min. RNA eluates were extracted with phenol/chloroform (1:1, pH 4.5) and chloroform and then ethanol precipitated in the presence of glycogen and resuspended in distilled water.

**RNA Ligation Assay—**RNA half-molecules were annealed for ligation in 2 x ligation buffer (1 x 40 mM HEPES, pH 7.8, 3 mM MgCl<sub>2</sub>, 2 mM spermidine, 5% glycerol) by heating to 85°C and slowly cooling to 70°C. The ligase reaction was also studied using purified enzyme and was found to be reversible in vitro.
E. coli 2′-5′ RNA Ligase Requires Base Modifications in tRNA Substrates—In previous studies of bacterial RNA ligase, activity had been assayed using substrates derived from tRNA precursors (pre-tRNAs) transcribed either in vivo in yeast or by endogenous RNA polymerase III activity in an S. cerevisiae extract (8). It was subsequently reported that pre-tRNA Phe transcripts from an artificial tRNA Phe gene by T7 RNA polymerase is an excellent substrate for the S. cerevisiae tRNA splicing enzymes (18), and this was also tested for ligation by E. coli RNA ligase. tRNA Phe half-molecules produced by digestion of T7-transcribed pre-tRNA Phe with S. cerevisiae tRNA splicing endonuclease were, however, poorly ligated in E. coli extracts (data not shown). Since tRNA Phe half-molecule substrates produced in vivo in yeast had previously been demonstrated to be ligated in E. coli extract (8), a requirement for some modification of substrate transcripts was implied.

In order to test the putative modification requirement, an S. cerevisiae tRNA Tyr gene was cloned under the control of a T7 RNA polymerase promoter. tRNA Tyr half-molecules derived from unmodified transcripts of this gene were poor substrates for the E. coli ligase, although they were utilized effectively by the S. cerevisiae tRNA splicing ligase. Fig. 1A shows that when pre-tRNA Tyr transcripts (lane M) were incubated in the same S. cerevisiae extracts utilized for RNA polymerase III transcription, tRNA Tyr half-molecules derived from them (lane 1) became substrates for the E. coli RNA ligase (lane 4). E. coli RNA ligase ligation of half-molecules derived from pre-tRNA Tyr modified in vitro produced a product of equivalent size to that produced by S. cerevisiae tRNA ligase or T4 RNA ligase and polynucleotide kinase (PNK) (lanes 2–3), although the bacterial enzyme is less efficient at joining these substrates. tRNA half-molecule substrates derived from modified T7 RNA polymerase transcripts were utilized by the E. coli RNA ligase as efficiently as yeast polymerase III transcripts, as quantified in a substrate titration experiment shown in Fig. 1B. This was presumably due to the formation of modified nucleosides in these transcripts, as observed by two-dimensional thin layer chromatography (TLC) of nuclease digests of substrate transcripts (data not shown). The identity and location of these modifications were not investigated further. The presence of a 2′-5′ linkage in ligated tRNA products was also observed by nuclease digestion and TLC (not shown), confirming that the reaction had been catalyzed by the previously described E. coli 2′-5′ ligase activity.

Pre-tRNA Phe T7 RNA polymerase transcripts modified by incubation in yeast extract were also used to produce substrates for ligation by the E. coli enzyme, although modified tRNA Tyr half-molecules were preferred as substrates by a factor of 2–3-fold over modified tRNA Phe half-molecules (Fig. 1C). Because tRNA Tyr half-molecules produced by endonuclease digestion of T7-transcribed, yeast extract-modified pre-tRNA Tyr were the most active substrates tested for ligation by the E. coli RNA ligase, they were chosen as substrates for quantitative assays of ligase activity during subsequent procedures.

Purification of the Ligase—A protocol for purification of the E. coli RNA ligase was developed and followed as detailed under “Experimental Procedures.” A quantitative profile of RNA ligase activity throughout the purification procedure is given in Table I. The complexity of the protein population at each purification step was assayed by SDS-PAGE and silver staining and is shown in Fig. 2. Briefly, cells were disrupted by sonication, and the lysate (Fig. 2, lane 1) was subjected to
Purification and Cloning of E. coli 2′-5′ RNA Ligase

TABLE I
Purification profile for preparation of E. coli 2′-5′ RNA ligase

| Fraction                  | Specific activity (fmol ligated/mg protein/min) | Purification factor | Yield |
|---------------------------|-----------------------------------------------|--------------------|-------|
| Crude extract             | 17.5                                         | 1                  | (77%) |
| S-100 Supernatant         | 22.8                                         | 1.3                | 100%  |
| DEAE effluent             | 246.5                                        | 14                 | ND    |
| Cellulose phosphate pool  | 1517                                         | 87                 | 60%   |
| Heparin Hyper-D pool      | 2381                                         | 136                | 37%   |
| E. coli tRNA-Sepharose    | 2780                                         | 159                | 23%   |
| Heparin Hyper-D (II) pool| 5826                                         | 335                | ND    |
| Superdex 75               | 8606                                         | 492                | 23%   |
| S. cerevisiae tRNA-Sepharose | 18930                                     | 1082               | 17%   |

FIG. 2. Purification of the E. coli RNA ligase enzyme. Silver-stained 16% SDS-PAGE gel of aliquots from individual purification steps. Samples were normalized according to total fraction volume. Lane 1, crude extract; lane 2, S-100 supernatant; lane 3, DEAE effluent; lane 4, cellulose phosphate pool; lane 5, heparin hyper-D pool; lane 6, E. coli tRNA-Sepharose pool; lane 7, heparin II pool; lane 8, Superdex 75 pool; lane 9, S. cerevisiae tRNA-Sepharose pool. The 20-kDa ligase protein (identified as described under “Reconstitution of Ligase Activity”) is indicated on the right.

Centrifugation at 100,000 × g. The S-100 supernatant (lane 2) was mixed batchwise with a DEAE anion-exchange resin, and the unbound fraction (lane 3) was loaded onto a cellulose phosphate column. RNA ligase activity was eluted from the cellulose phosphate by an increasing gradient of KCl, and fractions containing peak levels of RNA ligase activity were pooled for further purification by sequential binding to, and salt gradient elution from, heparin and E. coli tRNA affinity matrices. Pooled active fractions were subjected to gel filtration through Superdex 75 media. Peak RNA ligase activity fractions after gel filtration contained a mixture of at least six polypeptides (Fig. 2, lane 8). Superdex 75 peak fractions were then pooled for binding to a column of S. cerevisiae tRNA-Sepharose. RNA ligase activity eluted from this matrix with a salt gradient still contained several polypeptide species (lane 9).

Reconstitution of Ligase Activity from a Single Polypeptide Following SDS-PAGE—Since the optimized purification procedure did not yield a single homogeneous polypeptide, separation by denaturing electrophoresis was also utilized. A large scale purification was undertaken using 4 kg of E. coli cells as starting material. Final peak ligase activity fractions were pooled and concentrated by dialysis, and an aliquot of the concentrate was subjected to SDS-PAGE. Successive regions of the SDS-PAGE gel lane were excised and individually extracted. Eluted protein from each gel slice was precipitated and resuspended in guanidine HCl. After dialysis, eluates were assayed for RNA ligase activity (Fig. 3A) and examined by SDS-PAGE (Fig. 3B). Only a single eluate showed significant reconstituted activity (Fig. 3A, lane 5). The level of RNA ligase activity in this eluate was not affected by mixing with eluates of other gel slices (not shown). Fig. 3B shows that the active eluate contains only a single E. coli protein (of approximately 20 kDa) in addition to the bovine serum albumin carrier protein added during extraction (compare lane 5 to lane 2). A small amount of a polypeptide of about 20 kDa is present in the low molecular weight eluate (Fig. 3B, lane 3) which may explain the trace ligase activity observed in this fraction (Fig. 3A, lane 3). Due to its ability to reconstitute RNA ligase activity, the 20-kDa protein alone was presumed to be the E. coli RNA ligase.

Ligase Protein Sequencing and Identification of the RNA Ligase Gene—A second aliquot of the concentrated RNA ligase activity pool was subjected to SDS-PAGE and transferred to a nylon membrane, from which individual protein bands were excised for sequencing. 15 residues of amino-terminal sequence of the 20-kDa RNA ligase protein were obtained and are given in Fig. 4A. This sequence did not match any known protein or predicted ORF in the then current Genbank/EMBL data bases. A set of degenerate oligonucleotide primers corresponding to possible coding sequences for an internal segment of the amino-terminal sequence (shown in Fig. 4A) was synthesized for use in PCR. Fig. 4B shows that these oligonucleotides successfully amplified a DNA fragment of approximately 55 bp from E. coli cells or genomic DNA (lanes 1 and 2) but produced no product using S. cerevisiae cells as a template, or in reactions lacking either primers or template (lanes 3–5). The amplified DNA was cloned and sequenced.

In order to facilitate the cloning of the E. coli RNA ligase gene, its chromosomal location was determined using a genomic mapping blot. A radiolabeled DNA probe was created by PCR using the degenerate primers described above and hybridized to a membrane containing the Kohara “mini-set” of ordered, overlapping E. coli genomic phage clones (24). Hybrid-
BLAST algorithm (29) identified two highly similar protein sequences, predicted ORFs of unknown function from Methanococcus jannaschii (30) and Bacillus stearothermophilus (31). The E. coli, M. jannaschii, and B. stearothermophilus sequences are at least 23% identical and almost 50% similar over the entire length of the E. coli ligase protein (Fig. 5B). The Bacillus protein bears an extension of 129 amino acids at its carboxyl terminus, but this additional sequence does not have any significant matches in the available sequence data bases. Alignment of the three sequences reveals three highly conserved regions (a, b, and c in Fig. 5B) which may represent important functional domains of these proteins.

**Genomic Disruption of the RNA Ligase Gene**—A genomic disruption of the putative RNA ligase gene was performed in order to confirm that the protein identified by purification of RNA ligase activity was in fact the genuine ligase enzyme and to observe any phenotypes caused by a lack of RNA ligase function. The disruption was performed according to the method of Hamilton et al. (22) utilizing homologous recombination and subsequent resolution of a temperature-sensitive (ts) plasmid bearing an interrupted copy of the RNA ligase gene. First, the ligase ORF and 200 bp of upstream flanking sequence were amplified from genomic DNA using unique oligonucleotide primers and cloned into the pBlueScript vector to create the plasmid pBS-lig. A kanamycin resistance gene cassette was inserted into a unique restriction site at +45 bp in the ligase ORF, and the entire interrupted gene was subcloned into a plasmid containing a ts replicon and chloramphenicol resistance to create the plasmid pTSIL. After plasmid integration and resolution in a suitable RecA E. coli parent strain (as described under "Experimental Procedures"), 64 candidate colonies were recovered. One of these was found to be both kanamycin-resistant and chloramphenicol-sensitive at 43 °C, indicating stable chromosomal insertion. Insertion of the kanamycin cassette into the correct chromosomal locus was confirmed by PCR amplification of genomic DNA from individual colonies of the knockout isolate (Fig. 6A). Amplification using primers hybridizing either 1 bp upstream and at the 3’ end of the ORF (set a) or at the 5’ end of the ORF and 1.5 kb downstream (set b) gave a product which in the disrupted isolates was increased by 1200 bp, precisely the size of this cassette. When whole cell extracts of isolates of this knockout strain were assayed for RNA ligase activity, none was detected (Fig. 6B, compare lanes 3–5 with 6 and 7).

**Ligase Knockout Growth**—Effects of the disruption of RNA ligase expression in E. coli on overall fitness were examined by assaying bacterial growth under a variety of conditions. RNA ligase knockouts were viable and showed wild-type growth rates at temperatures ranging from 23 to 43 °C (not shown). The growth curve of knockout isolates at 37 °C was essentially identical to that of the parent strain (Fig. 7A). The effects of moderate amounts of extra chromosomal expression of the 2′-5′ RNA ligase in knockout and wild-type strains were also tested. E. coli strains were transformed with the pBS-lig construct which fortuitously supported expression of RNA ligase activity at approximately 10 times wild-type levels (as assayed by measuring specific activity in whole cell extracts) but still at a level below that detectable in crude extracts by SDS-PAGE and silver staining (not shown). All RNA ligase overproducing strains (Fig. 7B, B + D), but not those transformed with vector alone (A + C), were temperature-sensitive, being viable at 37 °C but unable to grow at 43 °C. These overproducing strains also showed a slow growth rate and reduced carrying capacity at stationary phase at 37 °C, as shown in Fig. 7A. Thus, the overproduction of E. coli RNA ligase has a toxic effect.

**Enzyme Equilibrium**—The equilibrium of the RNA ligation
reaction was studied in vitro in order to gain insight into the in vivo function of this enzyme. The time course of action of purified E. coli RNA ligase on tRNA\textsubscript{Tyr} half-molecules and on tRNA\textsubscript{Tyr} produced by ligation of tRNA half-molecules with E. coli RNA ligase (creating a 2'-5' linkage in the anticodon loop) was assayed. tRNA\textsubscript{Tyr} produced by ligation of half-molecules using T4 RNA ligase and PNK (to produce a 3'-5' junction) was also tested as a control. Fig. 8A shows that purified E. coli RNA ligase specifically cleaved 2'-5' linked substrates to fragments comigrating with authentic Tyr half-molecules (lanes 0-5, E. coli lig.) with approximately the same kinetics as ligation of half-molecules by that enzyme (lanes 0-5, Tyr 1/2's). tRNA\textsubscript{Tyr} with a 3'-5' linkage at the ligation junction was not cleaved (lanes 0-5, T4lig.). The identity of the tRNA cleavage products was confirmed by the ability of purified E. coli RNA ligase to rejoin them, as demonstrated in Fig. 8A (E. coli dig.). Beginning with either pure half-molecules or ligated tRNA, at 5 min of incubation the molar ratio of substrates to products approached the same value (4-5:1, halves:full-length), as quantified in Fig. 8C.

**DISCUSSION**

*E. coli* RNA Ligase Substrates—The results shown in Fig. 1 demonstrate that the E. coli 2'-5' RNA ligase requires modified nucleosides in artificial ligation substrates. This suggests that the in vivo substrate(s) of this enzyme is modified and therefore is likely to be a stable RNA as these modifications occur exclusively in stable RNAs in E. coli (32). Modified nucleosides may be recognized directly by the enzyme, as has been shown to occur in the interactions between some tRNAs and tRNA aminoacyl synthetases (33), or these base modifications may be required to stabilize the tRNA splicing substrates in a conformation that can be recognized by this enzyme. Modified nucleosides have been demonstrated to stabilize biologically active conformers of tRNAs in other systems (34, 35). The apparent requirement of *E. coli* 2'-5' ligase for nucleoside modifications...
and the preference shown by this enzyme for a subset of *S. cerevisiae* tRNA splicing substrates suggest that the *E. coli* ligase is likely to act upon a tRNA or tRNA-like molecule in vivo. Comparison of the four *S. cerevisiae* tRNA species that are ligated by the *E. coli* RNA ligase to the six that are not (8) does not reveal any obvious consensus of sequence or base modifications that might be recognized. The preference of the *E. coli* RNA ligase for yeast tRNA^Tyr^ half-molecules over tRNA^Phe^ half-molecules (Fig. 1C) suggests that this enzyme has the ability to discriminate among individual tRNA species.

**Ligase Purification**—Purification of the *E. coli* 2'-5' RNA ligase over 1000-fold from crude extracts provided highly purified protein fractions but not a single homogeneous polypeptide. Contaminating proteins remaining at the final stages of purification probably represent molecules with properties very similar to the 2'-5' RNA ligase, but are not likely to be components of a macromolecular complex as their concentrations peak in different fractions during gel filtration (data not shown). The tight binding of *E. coli* RNA ligase to immobilized prokaryotic and eukaryotic tRNA provides additional evidence that the ligase recognizes a tRNA or tRNA-like substrate in vivo. The ability of the RNA ligase protein to refold and reconstitute enzymatic activity after SDS-PAGE suggests a stable, self-folding structure for this protein.

**Ligase Gene Sequence**—Theoretical translation of the nucleotide sequence of the RNA ligase gene predicts a polypeptide with size, charge, and other biochemical properties in excellent agreement with those observed for the ligase protein. The ligase does not appear to be expressed as part of a multicistronic operon as the nearest upstream ORF with the same polarity is located about 12 kbp away, and the adjacent downstream ORF (encoding the sfiA protein) has its own promoter and regulatory elements (28). The cis-acting sequences controlling RNA ligase expression therefore remain to be determined.

The apparent conservation of the RNA ligase protein sequence between such distantly related bacterial species as *E. coli* and *B. stearothermophilus* and across kingdoms to the archaeote *M. jannaschii* suggests an ancient origin for this enzyme. This is in agreement with the observation of RNA ligase activity in extracts of a wide variety of bacterial species (8). The short blocks of high similarity between these predicted proteins as well as the lower dispersed similarity throughout suggest that the alignment is meaningful and is likely to represent a homologous origin and function for these proteins. If this alignment truly means that the 2'-5' RNA ligase is highly conserved between proteobacteria and Archaea, an important metabolic function for this enzyme is implied.

The RNA ligase enzyme may have been lost from some evolutionary branches between the low G + C Gram-positive bacteria and the Gamma division proteobacteria however, since RNA ligase activity was not originally detected in *M. jannaschii* (13). *Acinetobacter* species, although closely related to *E. coli*, do not harbor RNA ligase activity (29, 30). *Paracoccus denitrificans* and *Rhodopseudomonas viridis* are examples of divergent evolutionary branches in the Gammaproteobacteria, and some of these species may have lost the enzyme. The RNA ligase knockout in *E. coli* does not affect viability, growth rate, or chromosome stability (21, 22, 23, 24).

The RNA ligase protein sequence is conserved between the low G + C Gram-positive and Gamma division proteobacteria. *Escherichia coli* and *Bacillus stearothermophilus* ligases are approximately 53% identical in the alignment and may have diverged from a common ancestor between 0.45 and 87 million years ago (13). *Methanococcus jannaschii* and *E. coli* ligases are 26% identical and may have diverged from a common ancestor between 0.12 and 1.0 billion years ago (13).

**FIG. 6. Genomic knockout of the RNA ligase gene.** A, PCR amplification of ligase gene from various genomic DNA preparations. Products were separated on a 1.5% agarose gel and stained with ethidium bromide. M, DNA molecular mass standards; lanes 1 and 5, HB101; lanes 2 and 6, HS947 (knockout parent strain); lanes 3, 4, 7, and 8, disrupted isolates. Set a, primers located 1-kbp upstream and at the 3' end of the ligase ORF; set B, primers located at the 5' end of the ligase ORF and 1.5-kbp downstream. B, RNA ligation assay of lysates of indicated strains. Ligation reactions including 500 fmol were incubated at 30 °C for 10 min. Reactions supplemented with 1 μl each of the following: lane 1, buffer; lane 2, purified *E. coli* RNA ligase (Superdex 75 pool); lanes 3–7, crude lysates of HB101 (lane 3), HS947 (lane 4), a resolved cointegrate isolate not yet cured of the ts plasmid (lane 5), disrupted, resolved isolates cured of plasmid (lanes 6 and 7).

**FIG. 7. Growth phenotypes of RNA ligase knockout and overexpressing strains.** A, growth curves of *E. coli* strains at 37 °C derived by viable plating of three replicate cultures of each. Shown are the knockout parent strain HS947/RecA (○), HS947/RecA harboring the RNA ligase expression plasmid pBS-lig (▴), an RNA ligase knockout isolate (●), and a knockout harboring pBS-lig. (▲). Error bars represent one standard deviation. B, growth of strains at indicated temperatures on LB agar. A, HS947/RecA; B, HS947/RecA + pBS-lig; C, RNA ligase knockout; D, RNA ligase knockout + pBS-lig.
FIG. 8. Equilibrium of the E. coli RNA ligase. A, time course of forward and reverse reactions of highly purified E. coli RNA ligase. 4 μl of the Superdex 75 activity pool were added to 16 μl of the reactions containing the substrates indicated: Tyr 1/2’s, half-molecules derived from T7-transcribed, yeast-modified pre-tRNA⁹⁵ (4 pmol), E. coli lig₄tRNA⁹⁵, tRNA produced by ligation of tRNA⁹⁵ half-molecules with purified E. coli 2′-5′ RNA ligase (0.8 pmol); T4 lig₄tRNA⁹⁵, tRNA produced by ligation of tRNA⁹⁵ half-molecules with T4 RNA ligase and PNK in the presence of ATP (0.8 pmol). 4-μl aliquots were removed and stopped at the indicated times. B, quantification of the reactions in A, graphed according to the molar ratio of full-length tRNA to tRNA half-molecules at each time point. Shown are tRNA⁹⁵ half-molecules (●), E. coli-ligated tRNA⁹⁵ (▲), and T4-ligated tRNA⁹⁵ (○) substrates.

drastic decrease in size of its genome. This would imply (as did the lack of detectable activity in several bacterial species) that the RNA ligase may not perform a function absolutely necessary for bacterial survival but may be conditionally required under growth conditions encountered by a wide variety of species.

Genomic Disruption—Disruption of the genomic locus encoding the putative ligase protein confirmed that the correct polypeptide had been purified, since a complete loss of ligase activity in cell extracts ensued. This protein therefore appears to be the only enzyme in E. coli capable of ligating yeast tRNA half-molecules. The fact that genomic ligase knockout isolates were viable demonstrates that RNA ligase is not absolutely required for survival under laboratory growth conditions. Although the disrupted strains do not display a lethal phenotype, they can be examined for more subtle effects on growth and RNA metabolism. The availability of viable knockouts will also provide a useful null background for the expression of affinity-tagged or mutagenized ligase protein for use in further biochemical experiments. The toxic effects of moderate levels of ligase overexpression also imply some sort of interaction between the ligase enzyme and other cellular factors required for growth.

Examination of Enzyme Equilibrium—The reaction catalyzed by the E. coli ligase enzyme was shown to be fully reversible with an apparent equilibrium constant near unity, but favoring cleavage of 2′-5′ bonds. The tendency toward cleavage may perhaps be explained by the thermodynamics of phosphodiester bond cleavage and formation, which have been investigated thoroughly in the hammerhead ribozyme system. The favorable entropy of bond cleavage causes the internal equilibrium of the hammerhead ribozyme to favor cleavage of 3′-5′ phosphodiester bands to 2′,3′-cyclic phosphate and 5′-hydroxyl termini despite the unfavorable enthalpy associated with cyclic phosphate formation. For the cleavage reaction catalyzed by the E. coli RNA ligase, an increase of entropy in what is essentially a unimolecular reaction (given the tight structural association of tRNA half-molecules) may be due to the additional degrees of freedom available to released termini and to the disruption of water ordered in and about the closed, structured anticodon loop.

This observed equilibrium of cleavage and ligation appears to explain the maximum extent of ligation observed in in vitro activity assays. However, it begs the question of whether the function of the enzyme is to catalyze ligation or cleavage in vivo. Despite the fact that the equilibrium observed in vitro favors cleavage, the direction of the equilibrium in vivo will depend on the effective concentrations of substrates available for each reaction. If the true in vivo substrate is tRNA, ligated tRNA products are utilized for translation and thereby removed from the pool of substrates available to the ligase, then the ligation reaction will be favored. If, however, the cleavage products are removed by some process such as ribonucleolytic degradation, then the equilibrium will favor cleavage. To propose a cleavage function for the ligase enzyme in vivo, however, it is necessary to posit a source of substrates with 2′-5′ bonds, presumably in the context of a tRNA. No other E. coli enzyme is known or proposed that might form internal 2′-5′ linkages in a tRNA structure. A variety of endoribonuclease activities, however, could theoretically produce substrates for ligation by the E. coli RNA ligase, and in fact an activity
capable of doing so has also been observed in *E. coli* extracts.\(^2\) An enzyme capable of cyclizing free 3'-phosphates to 2',3'-cyclic phosphates has also recently been discovered in *E. coli*.\(^4\) Thus the available evidence, while circumstantial, favors a ligation function for this enzyme in vivo.

**Acknowledgments**—We gratefully acknowledge the contribution of unpublished genomic sequences by Drs. Keiko Takemoto and Hirotada Mori. This research was greatly assisted by scientific advice and guidance from Drs. Reinhard Rauhut, Jaime Arenas, Jeff Sampson, Peggy Saks, and Olke Uhlenbeck. Thanks to Dr. John Wagner and Christopher Trotta for critical reading of this manuscript and to Herbert J. Botts for his inspiration.

**REFERENCES**

1. Amitsur, M., Levitz, L., and Kaufmann, G. (1987) *EMBO J.* 6, 2499–2503
2. Peebles, C., Gegenheimer, P., and Abelson, J. (1983) *Cell* 32, 525–536
3. Greer, C., Peebles, C. L., Gegenheimer, P., and Abelson, J. (1983) *Cell* 32, 537–546
4. Gegenheimer, P., Gahius, H.-J., Peebles, C. L., and Abelson, J. (1983) *J. Biol. Chem.* 258, 8365–8373
5. Rusche, L., Pillar, K. J., and Sollner-Webb, B. (1995) *Mol. Cell. Biol.* 15, 2933–2941
6. Filipowicz, W., and Shatkin, A. J. (1983) *Cell* 32, 547–557
7. Kjems, J., Jensen, J., Olesen, T., and Garrett, R. A. (1989) *Can. J. Microbiol.* 35, 210–214
8. Greer, C., Javor, B., and Abelson, J. (1983) *Cell* 33, 899–906
9. Biniszkiewicz, D., Cesnaviciene, E., and Shub, D. A. (1994) *EMBO J.* 13, 4629–4635
10. Keindlhuber, B., and Shub, D. A. (1992) *Nature* 357, 173–176
11. Komine, Y., Adachi, T., Inokuchi, H., and Ozeki, H. (1990) *J. Mol. Biol.* 212, 579–598
12. Andachi, Y., Yamao, F., Muto, A., and Osawa, S. (1989) *J. Mol. Biol.* 209, 37–54
13. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., Mckenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Coton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, P. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fahrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science* 269, 496–512
14. Trujillo, M., Roux, D., Fueri, J. P., Cailla, H. L., and Rickenberg, H. V. (1987) *Eur. J. Biochem.* 169, 167–173
15. Evans, C., and Engelke, D. R. (1990) *Methods Enzymol.* 181, 439–450
16. Wallace, R., Johnson, P. F., Tanaka, S., Schold, M., Itakura, K., and Abelson, J. (1986) *Science* 209, 1396–1400
17. Sampson, J., and Saks, M. E. (1993) *Nucleic Acids Res.* 21, 4467–4475
18. Reyes, V., and Abelson, J. (1987) *Annu. Rev. Biochem.* 56, 90–106
19. Rauhut, R., Green, P. R., and Abelson, J. (1990) *J. Biol. Chem.* 265, 18160–18184
20. Schagger, H., Link, T. A., Engel, W. D., and von Jagow, G. (1986) *Methods Enzymol.* 129, 224–237
21. Wray, W., Beullink, T., Wray, V. P., and Hancock, R. (1981) *Anal. Biochem.* 118, 197–203
22. Hamilton, C. A., Aldea, M., Washburn, R. K., Babitzke, P., and Kushner, S. R. (1989) *J. Bacteriol.* 171, 4617–4622
23. Cornet, F., Mortier, I., Patte, J., and Louarn, J. M. (1994) *J. Bacteriol.* 176, 3188–3195
24. Kohara, Y. (1990) in *The Bacterial Chromosome* (Drlica, A., and Riley, M., ed) pp. 29–42, American Society for Microbiology, Washington, D. C.
25. Fujita, N., Mori, H., Yura, T., and Ishuhama, A. (1994) *Nucleic Acids Res.* 22, 1637–1639
26. Vicente, M., Kushner, S. R., Garrido, T., and Aldea, M. (1991) *Mol. Microbiol.* 5, 2085–2091
27. Loewen, P., and Henne-Aronis, R. (1994) *Annu. Rev. Microbiol.* 48, 53–80
28. Kawamukai, M., Utsumi, R., Takeda, K., Higashi, A., Matsuda, H., Choi, Y.-L., and Komano, T. (1991) *J. Bacteriol.* 173, 2644–2648
29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* 215, 403–410
30. Bult, C., White, O., Olsen, G. J. Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., Fitzgerald, L. M., Clayton, R. M., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Aams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A. (1996) *Science* 273, 1058–1073
31. Liong, E., and Ferenci, T. (1994) *Mol. Gen. Genet.* 243, 343–352
32. Bjork, G. (1987) in *Escherichia coli and Salmonella typhimurium*, *Cellular and Molecular Biology* (Neidhardt, F., ed) Vol. 1, pp. 719–731, American Society for Microbiology, Washington, D. C.
33. Perret, V., Garcia, A., Grosjean, H., Ebel, J.-P., Florentz, C., and Giege, R. (1996) *Nature* 344, 787–789
34. Arnez, J., Steitz, T. A. (1994) *Biochemistry* 33, 7560–7567
35. Prykorska, A. (1995) *Biochimie* (Paris) 77, 109–112
36. Hertel, K., and Uhlenbeck, O. C. (1995) *Biochemistry* 34, 1744–1749

---

\(^2\) E. Arn, unpublished observations.

\(^4\) W. Filipowicz, personal communication.