tRNA^{His} maturation: An essential yeast protein catalyzes addition of a guanine nucleotide to the 5′ end of tRNA^{His}

Weifeng Gu,[1] Jane E. Jackman,[1] Amanda J. Lohan,[2] Michael W. Gray,[2] and Eric M. Phizicky[1,3]

[1]Department of Biochemistry and Biophysics, University of Rochester School of Medicine, Rochester, New York 14642, USA; [2]Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax NS, B3H 1X5, Canada

All tRNA^{His} molecules are unusual in having an extra 5′ GMP residue (G_{-1}) that, in eukaryotes, is added after transcription and RNase P cleavage. Incorporation of this G_{-1} residue is a rare example of nucleotide addition occurring at an RNA 5′ end in a normal phosphodiester linkage. We show here that the essential *Saccharomyces cerevisiae* ORF YGR024c (*THG1*) is responsible for this guanylyltransferase reaction. Thg1p was identified by survey of a genomic collection of yeast GST-ORF fusion proteins for addition of [α-^{32}P]GTP to tRNA^{His}. End analysis confirms the presence of G_{-1}. Thg1p is required for tRNA^{His} guanylylation in vivo, because cells depleted of Thg1p lack G_{-1} in their tRNA^{His}. His$_6$-Thg1p purified from *Escherichia coli* catalyzes the guanylyltransferase step of G_{-1} addition using a ppp-tRNA^{His} substrate, and appears to catalyze the activation step using p-tRNA^{His} and ATP. Thg1p is highly conserved in eukaryotes, where G_{-1} addition is necessary, and is not found in eubacteria, where G_{-1} is genome-encoded. Thus, Thg1p is the first member of a new family of enzymes that can catalyze phosphodiester bond formation at the 5′ end of RNAs, formally in a 3′–5′ direction. Surprisingly, despite its varied activities, Thg1p contains no recognizable catalytic or functional domains.

[*Keywords*: *Saccharomyces cerevisiae*; guanylyltransferase; tRNA processing; YGR024c]

Supplemental material is available at http://www.genesdev.org.

Received August 29, 2003; revised version accepted October 15, 2003.

All known tRNA^{His} molecules have an unusual 5′ end consisting of an additional GMP residue, called G_{-1} (Fig. 1A; Sprinzl et al. 1998; Marck and Grosjean 2002). In prokaryotes, this G_{-1} residue is genome-encoded, and originates from anomalous RNase P cleavage of the precursor tRNA^{His} (pre-tRNA^{His}) at the –1 position to generate tRNA with an additional G_{-1}:C_{73} base pair in the acceptor stem (Orellana et al. 1986; Burkard et al. 1988). However, in the cytoplasm of eukaryotes (Sprinzl et al. 1998), and in at least one mitochondrial species (L'Abbe et al. 1990), G_{-1} is not derived from tRNA^{His} gene sequence, but instead is added after transcription. Furthermore, in these cases the added G_{-1} is across from A_{73} in the tRNA (Sprinzl et al. 1998; Marck and Grosjean 2002).

The posttranscriptional addition of G_{-1} is intriguing for three reasons: First, it is one of two known reactions that result in extension of a polynucleotide in a 3′ to 5′ direction by formation of a normal phosphodiester bond. The other example of this type of reaction is found in the amoeboid protozoon *Acanthamoeba castellanii*, which contains an activity that edits a number of mitochondrial tRNAs by removal of up to three unpaired nucleotides at the 5′ end, followed by templated polymerization in the 3′–5′ direction (Lonergan and Gray 1993; Price and Gray 1999). A very similar tRNA editing activity appears to act in the mitochondria of several chytridiomycete fungi (Laforest et al. 1997; Bullerwell et al. 2003). Second, of more than 500 sequenced tRNAs from all kingdoms, only tRNA^{His} species and one exceptional tRNA^{Phe} species (Schnare et al. 1985) have an extra 5′ nucleotide. Third, the added G_{-1} is a critical determinant for aminoacylation of tRNA^{His} by the corresponding synthetase in yeast (Rudinger et al. 1994; Nameki et al. 1995). Thus, this modification is likely critical for cell function.

Because of its importance and unique biochemical character, we set out to identify the gene product(s) responsible for G_{-1} addition activity. Earlier work had shown the existence of a tRNA^{His} guanylyltransferase activity in extracts from *Drosophila* (Cooley et al. 1982) and yeast (Williams et al. 1990) that required both ATP and a guanine nucleotide. Subsequent work led to the identification and purification of a 58-kD polypeptide from yeast that was implicated in activity (Pande et al. 1991), and a proposed mechanism very similar to that

[3]Corresponding author.
E-MAIL eric_phizicky@urmc.rochester.edu; FAX (585) 271-2683.
Article published online ahead of print. Article and publication date are at http://www.genesdev.org/cgi/doi/10.1101/gad.1148603.

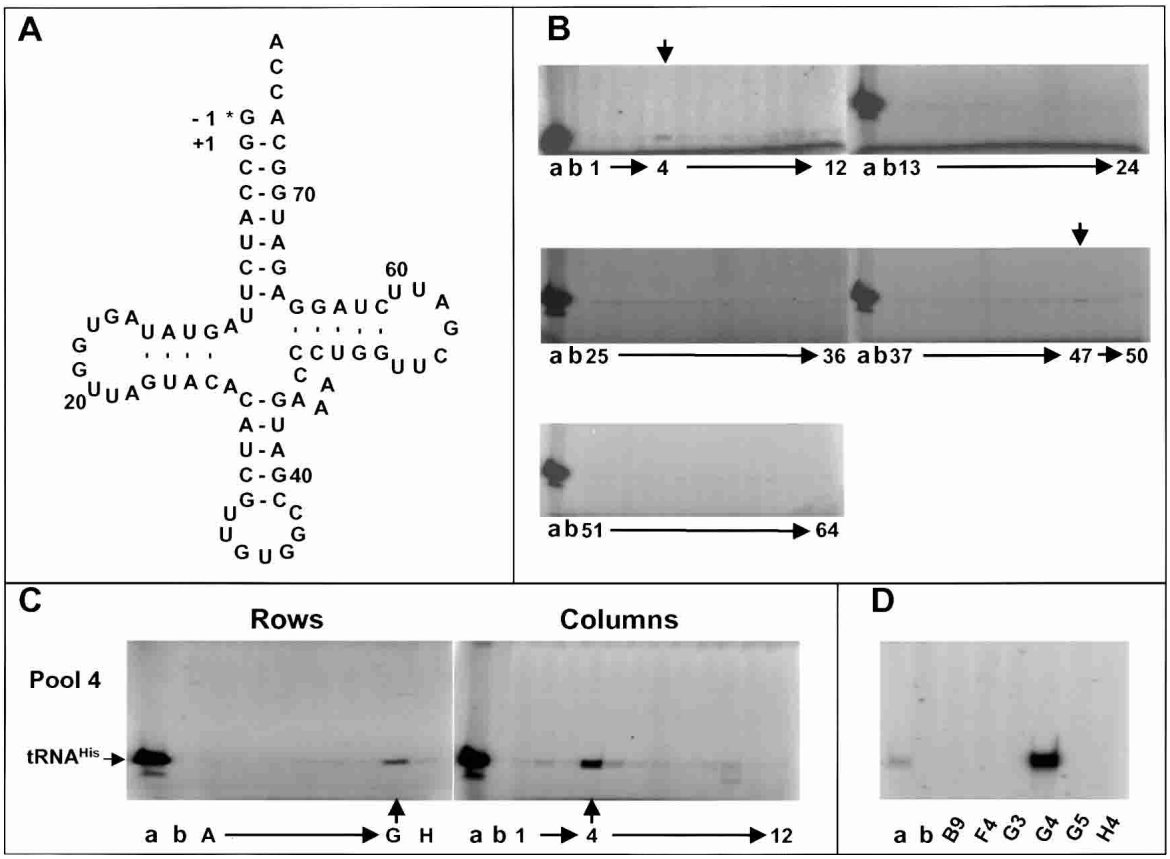

Figure 1. Identification of yeast ORFs associated with tRNA^{His} guanylyltransferase activity. (*A*) Cloverleaf structure of *S. cerevisiae* cytoplasmic tRNA^{His}. Numbers indicate standard nt numbering positions; +1, the first nt encoded by the genome; –1, the extra nt added posttranscriptionally. Note that the tRNA is 75-nt long before G_{-1} addition, because nt 17 and nt 47 are not present in this tRNA, and nt 20A is present. (*B*) Assay of a genomic collection of pools of purified yeast GST-ORF fusion proteins for tRNA^{His} guanylyltransferase activity. Substrate p-tRNA^{His} (0.4 µM) was incubated in 40-µL reaction mixtures containing 3 mM ATP, 0.6 µM [α-^{32}P]GTP, and ~0.01 µg/µL protein from 64 pools of purified GST-ORF fusion proteins, each derived from 96 yeast strains, and products were resolved on gels as described in Materials and Methods. *a*, S100 extract (2 µg/µL); *b*, no extract. (*C*) Assay of subpools from pool 4 for activity. Substrate p-tRNA^{His} was incubated with pools of GST-ORFs derived from the strains in rows *A-H* and columns *1–12* from plate 4, as indicated. *a*, S100 extract; *b*, no extract. (*D*) Assay of individual purified proteins from plate 4 for activity. *a*, S100 extract; *b*, no extract; B9, F4, G3, G4, G5, and H4 each correspond to the purified GST-ORF fusion protein from the corresponding position of microtiter plate 4.

employed by other RNA ligases, DNA ligases, and guanylyltransferases (Jahn and Pande 1991). Guanylyl transfer was accomplished by adenylylation of the enzyme, transfer of the AMP to the 5′-phosphate of tRNA^{His} in a 5′–5′ pyrophosphate linkage, and displacement of the AMP moiety by attack of the 3′-OH of the guanine nucleotide to effect G addition (Williams et al. 1990; Jahn and Pande 1991). No subsequent work has been reported on this activity.

Using a previously described biochemical genomics approach (Martzen et al. 1999), and an assay based on incorporation of [α-^{32}P]GTP into a tRNA^{His} substrate, we identified a 28-kD protein, encoded by the essential ORF YGR024c, that copurifies with tRNA^{His} guanylyltransferase activity. We have shown that this protein, which we designate Thg1p (tRNA^{His} guanylyltransferase), is responsible for tRNA^{His} guanylylation in vivo. Thg1p is highly conserved in Eucarya (eukaryotes) and Archaea (archaebacteria) but absent from Bacteria (eubacteria), consistent with the requirement for this activity in eukaryotes but not eubacteria. Remarkably, when expressed and purified from *Escherichia coli*, Thg1p has several of the biochemical activities required of tRNA^{His} guanylyltransferase, despite the fact that it bears little resemblance to known guanylyltransferases, ligases, or pyrophosphatases that catalyze similar types of reactions.

Results

Identification of two GST-ORF fusion proteins that copurify with tRNA^{His} guanylyltransferase activity

To identify the protein(s) responsible for tRNA^{His} guanylyltransferase activity, we monitored incorporation of [α-^{32}P]GTP into a synthetic 75-mer tRNA^{His} substrate

beginning at the +1 position, using polyacrylamide gel electrophoresis (PAGE). Based on previous reports we also added ATP to the reaction mixtures (Cooley et al. 1982; Williams et al. 1990; Pande et al. 1991). Because the authentic substrate for the guanylyltransferase activity is produced by the action of RNase P, which leaves a 5′-terminal phosphate, the substrate tRNA was transcribed in the presence of a 10-fold molar excess of GMP (pG) over GTP to produce tRNA^{His} substrate that initiates primarily with pG at the +1 position (we refer to this substrate as "p-tRNA^{His}"). The transcript also terminates with the sequence CCA, which may or may not be the authentic 3′ end of the in vivo substrate, depending on whether CCA addition occurs in vivo before or after the G addition reaction. Incubation of this tRNA substrate in a crude yeast S100 extract in the presence of [α-^{32}P]GTP and ATP, followed by gel electrophoresis, yields a prominent signal at the position of mature tRNA (Fig. 1B, lane a).

With this assay we identified two gene products that copurify with [α-^{32}P]GTP incorporation activity, employing a biochemical genomics approach (Martzen et al. 1999) that we used previously to assign several other gene products to tRNA modification activities (Alexandrov et al. 2002; Xing et al. 2002; Jackman et al. 2003). In this approach, activity is assayed in 64 pools of purified yeast GST-ORF fusion proteins derived from a genomic collection of 6144 yeast strains, each of which expresses a unique yeast GST-ORF fusion protein. Active pools are then deconvoluted by preparation and analysis of subpools of purified GST-ORF fusion proteins to assign the activity to a single strain and ORF.

As shown in Figure 1B, examination of the 64 pools for [α-^{32}P]GTP incorporation activity reveals significant labeled product tRNA in pool 4, and possibly in pool 47. To identify the ORF associated with activity in each pool, we analyzed subpools of GST-ORF fusion proteins derived from rows and columns of the corresponding microtiter plates of strains. As shown in Figure 1C, the activity in pool 4 is found in pools from row G and column 4 of the microtiter plate, implicating a strain that expresses YGR024c. To confirm this assignment, we purified and assayed GST-YGR024c fusion protein from this strain. As shown in Figure 1D, GST-YGR024c protein has substantial activity relative to that of control GST-ORF fusion proteins from other strains. Finally, we isolated and sequenced the plasmid from this strain, retransformed the plasmid into yeast, purified the expressed fusion protein, and confirmed that this preparation of GST-YGR024c has activity (see below). Similar experiments demonstrated that the activity in pool 47 is due to the GST fusion protein from ORF YDL076c (data not shown). ORF YGR024c encodes an essential 28-kD protein that was previously uncharacterized, and ORF YDL076c encodes a nonessential 33.8-kD protein.

Purified GST-YGR024c protein and GST-YDL076c protein have tRNA^{His} guanylyltransferase activity

The [α-^{32}P]GTP incorporation activities that copurify with GST-YGR024c and with GST-YDL076c exhibit three biochemical properties expected of tRNA^{His} guanylyltransferase:

First, as expected for tRNA^{His} guanylyltransferase activity, the incorporation of GTP into tRNA-sized material is specific for tRNA^{His}. Thus, labeled GTP is incorporated into tRNA^{His} (Fig. 2A, lane e; data not shown) but not into tRNA^{Phe} (Fig. 2A, lane b; data not shown).

Second, the product tRNA has the expected structure. Digestion of the purified product tRNA with phosphatase releases inorganic phosphate, demonstrating that the α-phosphate is exposed (Fig. 2B; data not shown). This result suggests that addition occurs at the 5′ end, because addition at the 3′ end would likely involve formation of a phosphodiester bond involving the α phosphate of GTP, which would make the bond phosphatase-resistant. The phosphatase sensitivity of the phosphate label in the product tRNA also rules out formation of a Gp*-p-tRNA adduct in which GMP is covalently attached to p-tRNA by a 5′–5′ phosphoanhydride bond, as would occur in the activation by a ligase, or in the second step of mRNA capping (Shuman and Schwer 1995; Ho and Shuman 2002). However, these results do not rule out some sort of exchange reaction in which the pG residue at the +1 position of the tRNA substrate is somehow exchanged with the labeled GTP. To prove directly that the tRNA product is one nucleotide longer than the substrate tRNA^{His}, we generated a partial RNase T1 cleavage ladder from the product tRNAs after reaction with GST-ORF fusion proteins, and compared the pattern with that formed from 5′ end-labeled transcripts of 76-mer p-tRNA^{His} and 75-mer p-tRNA^{His} substrate. As shown in Figure 2C, the RNase T1 cleavage pattern produced from reaction with GST-YGR024c (lane c) or with GST-YDL076c (lane d) exactly matches that produced from 5′ end-labeled 76-mer p-tRNA^{His} (lane a), is consistently one nucleotide longer than the pattern from 5′-end-labeled 75-mer p-tRNA^{His} substrate (lane b), and is completely different from that produced from 75-mer p-tRNA^{His} substrate labeled at its 3′ end with pCp (lane e). Thus the product of the reaction with tRNA^{His} substrate has an extra pG residue at its 5′ end.

Third, the G_{-1} residue is added in a normal phosphodiester linkage, as P1 nuclease treatment of the product yields mostly labeled pG, as well as some minor amount of material that comigrates with GTP (data not shown).

ORF YGR024c, but not ORF YDL076c, is required for tRNA^{His} guanylyltransferase activity in vivo

To determine whether ORFs YGR024c and/or YDL076c are required for tRNA^{His} guanylyltransferase activity, we analyzed the tRNA of strains that lacked each protein.

Because ORF YGR024c is essential, we first constructed a strain in which expression of the only copy of ORF YGR024c was under regulated control, and then analyzed tRNA^{His} from this strain when expression was repressed. This strain (WG18, with relevant genotype: *ygr024c*-Δ::kan^r, P_{GAL}-YGR024c) is healthy on medium

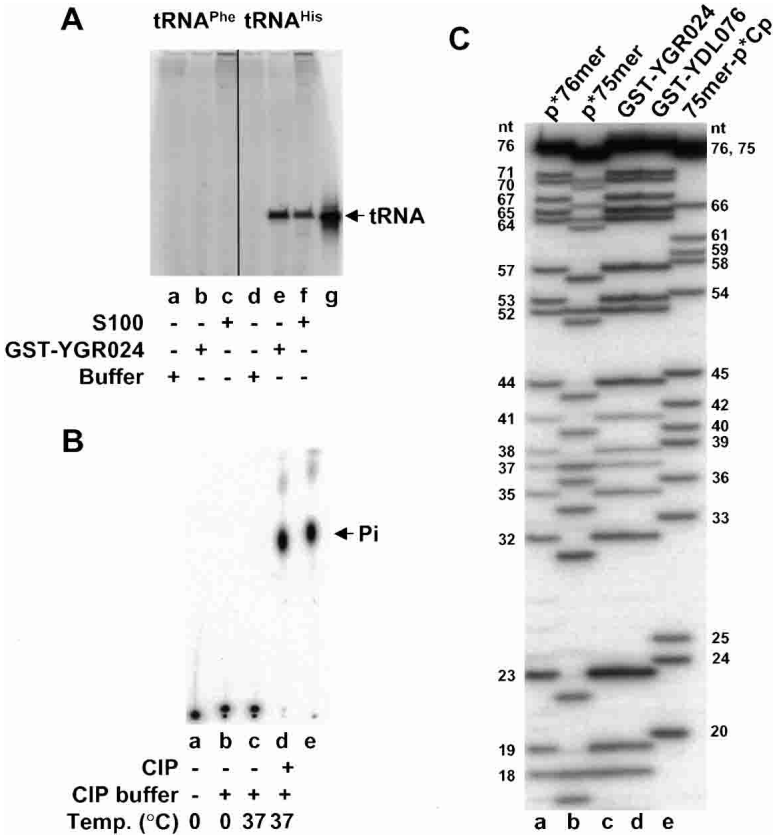

Figure 2. Analysis of the activity of purified GST-YGR024c fusion protein. (*A*) tRNA substrate specificity of YGR024p. Here, 0.4 µM p-tRNA^{Phe} (lanes *a–c*) or 0.4 µM p-tRNA^{His} (lanes *d–f*) was incubated in reaction mixtures containing 100 µM ATP, 0.6 µM [α-^{32}P]GTP, and GST dialysis buffer (lanes *a*,*d*), 0.5 µM GST-YGR024p (lanes *b*,*e*), or 2.4 µg/µL S100 extract (lanes *c*,*f*), and tRNA was resolved by PAGE as described in Materials and Methods. Lane *g*, [5′-^{32}P] tRNA^{His}. (*B*) Phosphatase treatment of guanylylated tRNA^{His}. Gel-purified tRNA product from reaction with GST-YGR024p was incubated with calf intestinal phosphatase (CIP), and products were resolved by thin layer chromatography as described in Materials and Methods. *a*, tRNA incubated on ice; *b*, tRNA on ice in CIP buffer; *c*, tRNA in CIP buffer at 37°C; *d*, tRNA incubated with CIP in CIP buffer at 37°C; *e*, ^{32}P-labeled inorganic phosphate. (*C*) Partial RNase T1 digestion of guanylylated tRNA^{His}. Gel-purified guanylylated tRNA^{His} and labeled tRNA standards were incubated with RNase T1 under partial digestion conditions, as described in Materials and Methods, and products were resolved by 12% PAGE. *a*, [5′-^{32}P]tRNA^{His} 76-mer, beginning with G_{-1} (p*76mer); *b*, [5′-^{32}P]tRNA^{His} 75-mer, beginning with G_{+1} (p*75mer); *c*,*d*, p-tRNA^{His} substrate guanylylated with GST-YGR024p and GST-YDL076p, respectively, in the presence of 100 µM ATP and [α-^{32}P]GTP; *e*, 75-mer tRNA^{His} 3′-labeled with *pCp. Numbers on the left and right indicate the length of partially digested fragments in lanes *a* and *e*, respectively.

containing galactose, and when shifted to medium containing glucose (repressing conditions), stops growing after two rounds of replica-plating. In liquid cultures, shift to glucose-containing medium results in progressively slower growth beginning about 17 h later (Fig. 3A). This gradual onset of slow growth is similar to previous observations with a strain expressing its only copy of tRNA ligase under P_{GAL} control (Phizicky et al. 1992).

Primer extension experiments demonstrate that WG18 cells conditionally lacking YGR024p accumulate tRNA ^{His} that is one nucleotide shorter at its 5′ end than that from wild-type cells (Fig. 3B, cf. lanes e,f and lanes g–j). This shorter tRNA^{His} predominates in WG18 cells by 20 h after the shift to glucose (Fig. 3B, lane i), about the time that growth begins to slow; in contrast, tRNA-^{His} in a control wild-type strain (WG12) is the normal length throughout the growth period (Fig. 3B, lanes e,f). We note the presence of a distinct tRNA^{His} primer extension pause at position +5 in tRNA from both wild-type strains and strains depleted of YGR024c; this pause is due to 2′-*O*-methyl adenosine at position +4 in tRNA^{His}. Because the 5′ end of a control tRNA, tRNA$^{Lys}_{UUU}$, is unaffected in the WG18 strain lacking YGR024c (Fig. 3C, cf. lanes e,f and g–j), we conclude that the shorter tRNA^{His} 5′ end observed in WG18 is specific for this tRNA. We note that the primer extension observed upstream of position +1 of tRNA^{Lys} corresponds exactly to the sequence predicted from the genome sequence (Fig. 3C, lanes a–d), and thus corresponds to pre-tRNA^{Lys}. Analysis of extracts from the WG18 strain conditionally lacking YGR024c shows that guanylyltransferase activity is decreased to undetectable levels after 20 h in glucose (Fig. 3D, lanes k,l), whereas extracts from the control strain have normal activity (Fig. 3D, lanes e,f). These observations prove that ORF YGR024c is required for the addition of G_{-1} to the 5′ end of tRNA^{His} in vivo, and we have therefore assigned the name *THG1* to ORF YGR024c.

In contrast, ORF YDL076c does not appear to be required for tRNA^{His} guanylyltransferase activity in vivo. We analyzed steady-state levels of tRNA^{His} from two independently constructed strains lacking this ORF, and each has full-length tRNA^{His} (data not shown). This suggests that ORF YDL076c does not affect tRNA^{His} production in vivo, although it could have a minor effect, or there may be a redundant cellular activity that can replace its function. We have not further analyzed the function of ORF YDL076c.

THG1 (ORF YGR024c) encodes a protein with tRNA^{His} guanylyltransferase activity

Because GST-Thg1p copurifies with yeast tRNA^{His} guanylyltransferase activity (Fig. 1), this implies either that Thg1p itself has tRNA^{His} guanylyltransferase activity, or that Thg1p is part of a complex with this activity. To distinguish between these possibilities, we purified yeast His$_6$-Thg1p expressed in *E. coli*, where it would be free of

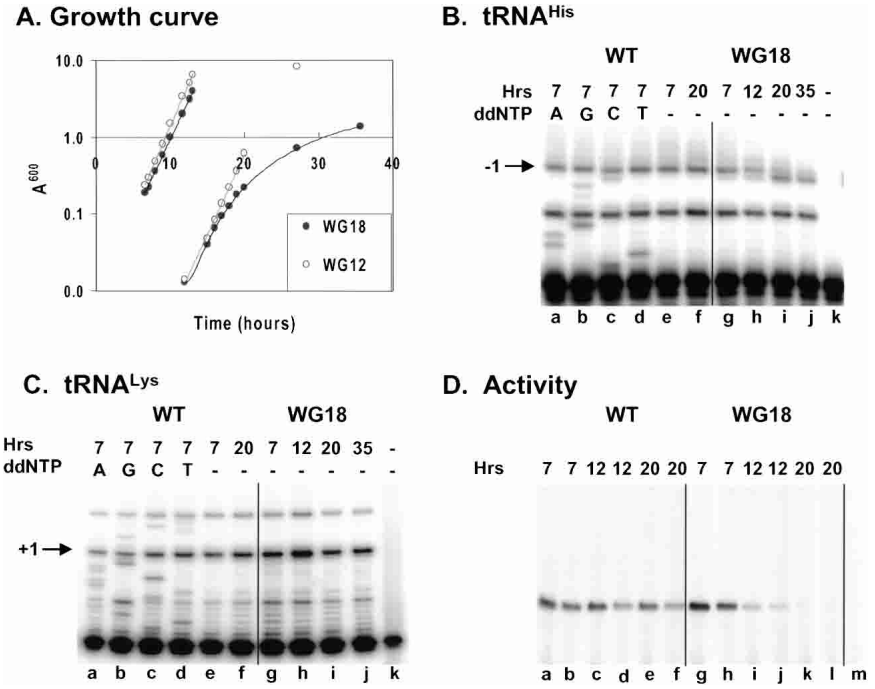

Figure 3. Effect of lack of ORF YGR024c protein on cell growth and on tRNA^{His} guanylylation. Strain WG18 conditionally lacking YGR024c (relevant genotype: *ygr024c*-Δ P_{GAL}-YGR024c) and the corresponding wild-type control strain (WT) WG12 were grown in rich medium containing galactose and inoculated in medium containing glucose. Growth, tRNA 5′-end maturation, and guanylyltransferase activity of extracts were monitored as a function of time. (*A*) Growth defect of strain conditionally lacking YGR024c. (●) Growth of strain conditionally lacking YGR024c. (○) Growth of WT cells. (*B*) Analysis of the 5′ end of tRNA^{His}. Bulk RNA was purified at each time point, and analyzed by primer extension as described in Materials and Methods to determine the 5′ ends of tRNA^{His}. Lanes *a–d*, sequencing reactions using RNA purified from WG12 at 7 h; *e,f*, RNA purified at 7 h (*e*) and 20 h (*f*) from WG12 after switch to glucose-containing medium; *g–j*, RNA purified from WG18 strain at various times after switch to glucose-containing medium, as indicated; *k*, primer alone. Arrow at –1 indicates position of the mature 5′ end of guanylylated tRNA^{His}. (*C*) Analysis of the 5′ end of tRNA^{Lys}. Bulk RNA was analyzed by primer extension to determine the 5′ end of tRNA^{Lys}, as described in *B* and Materials and Methods. (*D*) Assay for tRNA^{His} guanylyltransferase activity in crude extracts. Crude extracts made at different times after switch to glucose medium were assayed for tRNA^{His} guanylyltransferase activity as described in Materials and Methods. *a–f*, crude extracts from WT strain made at different time points, each assayed with 0.2 μg/μL (*a,c,e*) and 0.02 μg/μL (*b,d,f*) protein; g–*i*, crude extracts from WG18 strain made at different time points, assayed with 0.2 μg/μL (*g,i,k*) and 0.02 μg/μL (*h,j,l*) protein; *m*, no protein.

any yeast contaminants, to compare its activity with GST-Thg1p purified from yeast. As shown in Figure 4A, a strain expressing His$_6$-Thg1p expresses a prominent polypeptide of ~30 kD (lane a) that is absent in the control extract (lane d), and this polypeptide is readily purified by immobilized metal ion affinity chromatography (lanes b,c). The resulting protein has few visible contaminants when visualized by Coomassie staining, and is of comparable purity to GST-Thg1p obtained from yeast by glutathione agarose chromatography (Fig. 4B).

Two surprising results emerged from the analysis of the activity of recombinant His$_6$-Thg1p purified from *E. coli* and of GST-Thg1p purified from yeast, as shown in Figure 5A.

First, His$_6$-Thg1p is active in the absence of yeast components when assayed with our normal p-tRNA^{His} sub-

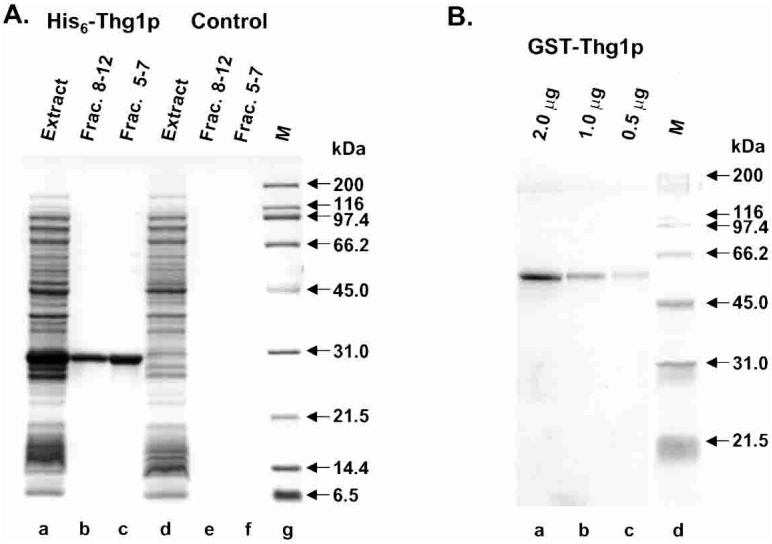

Figure 4. SDS-PAGE analysis of GST-Thg1p purified from yeast and His$_6$-Thg1p purified from *E. coli*. *THG1* (YGR024c) expressed in yeast as GST-Thg1p and in *E. coli* as His$_6$-Thg1p, was purified as described in Materials and Methods and analyzed by SDS-PAGE on an 8%–16% gel (BioRad). (*A*) Analysis of His$_6$-Thg1p purified from *E. coli* and mock purification. *a,d*, crude extracts from transformed *E. coli* cells and mock-transformed cells; *b,c*, purified His$_6$-Thg1p from pooled elution fractions 8–12 (3 μg) and 5–7 (6 μg); *e,f*, corresponding mock purification fractions; *g*, SDS-PAGE broad range protein standards (BioRad). Samples were visualized with Coomassie staining. (*B*) Analysis of GST-Thg1p purified from yeast. *a–c*, different amounts of purified GST-Thg1p as indicated; *d*, SDS-PAGE broad-range protein standards. Samples were visualized with silver staining.

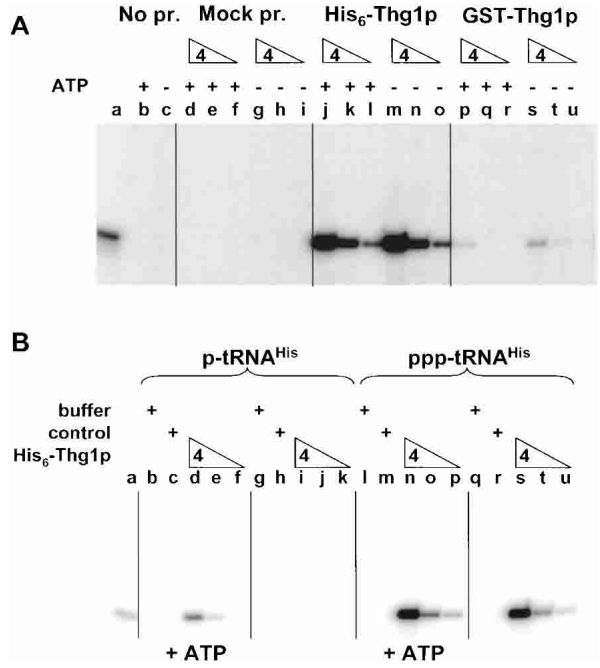

Figure 5. Analysis of guanylyltransferase activity of purified His_6-Thg1p and GST-Thg1p. (*A*) Analysis with standard $tRNA^{His}$ substrate. Purified proteins were assayed for guanylyltransferase activity in the presence or absence of 50 µM ATP as indicated, as described in Materials and Methods. *a*, [5′-^{32}P] p-$tRNA^{His}$ substrate; *b*,*c*, no protein; d–i, pooled fractions 8–12 of mock purification from *E. coli* (see Fig. 4*A*, lane *e*) in the presence (*d–f*) or absence (*g–i*) of ATP; *j–o*, pooled fractions 8–12 of purified His_6-Thg1p in the presence (*j–l*) or absence (*m–o*) of ATP; *p–u*, GST-Thg1p in the presence (*p–r*) or absence (s–u) of ATP. Triangles, fourfold serial dilutions of protein (0.5 µM, 0.125 µM, and 0.032 µM). (*B*) Analysis of His_6-Thg1p with defined p-tRNA and ppp-tRNA substrates. Defined p-$tRNA^{His}$ and ppp-$tRNA^{His}$ substrates, prepared as described in Materials and Methods, were assayed for activity with His_6-Thg1p in the presence or absence of 50 µM ATP, and products were resolved by PAGE as described above. *a*, [5′-^{32}P] $tRNA^{His}$ substrate; *b–k*, assay of p-$tRNA^{His}$ substrate in the presence (*b–f*) or absence (*g–k*) of ATP. *l–u*, assay of ppp-$tRNA^{His}$ substrate in the presence (*l–p*) or absence (*q–u*) of ATP; *b*,*g*,*l*,*q*, buffer controls; *c*,*h*,*m*,*r*, mock purification controls; *d–f*, *i–k*, *n–p*, *s–u*, fourfold serial titrations of His_6-Thg1p from 0.5 µM to 0.032 µM.

strate in the presence of [α^{32}-P]GTP (Fig. 5A, lanes j–l). This was unexpected because, as described below, we could detect no obvious similarity between Thg1p and any other protein that catalyzes activities expected for this type of reaction. Indeed, the preparation of His_6-Thg1p is substantially more active than that of GST-Thg1p (Fig. 5A, cf. lanes j–l and p–r). One possible explanation of this result is that activity of the yeast-derived protein is inhibited by the dimerization of the GST tag (Riley et al. 1996); however, it could also be the case that the yeast-derived protein copurifies with natural inhibitors, or that it is modified posttranslationally and has less activity in this assay.

Second, both yeast-derived and *E. coli*-derived Thg1p are comparably active in the presence or absence of ATP (Fig. 5A, cf. lanes j–l and m–o, and lanes p–r and s–u). This was unexpected, because addition of G_{-1} should require activation of the tRNA 5′ end by formation of a high-energy bond and, based on previous work, ATP was expected to be involved in the activation (Cooley et al. 1982; Williams et al. 1990; Jahn and Pande 1991; Pande et al. 1991).

There are three reasonable explanations for this result. First, GTP might activate the tRNA, possibly forming a Gp-p-tRNA intermediate. This interpretation would suggest that GTP is used both to activate the tRNA and to effect transfer of the guanine nucleotide to the 5′ end. Second, the product formed in the absence of ATP might be different than that found in the presence of ATP. However, experiments similar to those in Figure 2 demonstrate that the product formed with His_6-Thg1p and [α^{32}-P]GTP, in the presence or absence of ATP, has the expected structure: the tRNA has an extra guanine nucleotide on its 5′ end in normal 5′–3′ linkage, primarily as a GMP residue (data not shown). Third, the synthetic $tRNA^{His}$ substrate that is used for guanylylation in the absence of ATP could be the small fraction that begins with a tri-phosphorylated nucleotide (ppp-$tRNA^{His}$) instead of a mono-phosphorylated nucleotide (p-$tRNA^{His}$). This would fulfill the energetic requirements of the guanylyltransferase reaction.

To determine the exact substrate and ATP dependence of His_6-Thg1p, we separately prepared and assayed two substrates with different defined 5′ ends: a ppp-$tRNA^{His}$ substrate, made by transcription without inclusion of pG; and a p-$tRNA^{His}$ substrate, made by transcription of a $tRNA^{His}$ template with an 18-nt 5′-extension, followed by directed RNase H cleavage to generate $tRNA^{His}$ beginning with a 5′ monophosphate at G_{+1} (see Materials and Methods). From the data in Figure 5B, we draw two conclusions. First, Thg1p is active with the ppp-$tRNA^{His}$ substrate, in the presence or absence of ATP (Fig. 5B, lanes n–p, s–u). Second, Thg1p activity with the p-$tRNA^{His}$ substrate is undetectable in the absence of ATP (Fig. 5B, lanes i–k), but greatly stimulated by addition of ATP (Fig. 5B, lanes d–f). These results suggest that Thg1p is in fact the catalytic component for the guanylyl transfer step of the reaction, and that an energy source such as ATP is required to activate the p-$tRNA^{His}$ substrate for reaction with His_6-Thg1p. Based on the previously postulated mechanism of this reaction (Jahn and Pande 1991), the most likely activation step involves formation of an Ap-p-tRNA intermediate in which the AMP residue derived from ATP is in a 5′–5′ phosphoanhydride linkage.

We determined the nature of the activation step by incubation of His_6-Thg1p with p-$tRNA^{His}$ in the presence of [α-^{32}P]ATP, but in the absence of GTP. This yields labeled tRNA (data not shown; see also below), which was purified and further analyzed to determine its identity. As shown in Figure 6A, digestion of the labeled tRNA with P1 nuclease yields mostly material consistent with the expected adenylylated product Ap*-pG, bearing a 5′–5′ pyrophosphate linkage. As expected for Ap*-pG, it is resistant to phosphatase treatment (Fig. 6A,

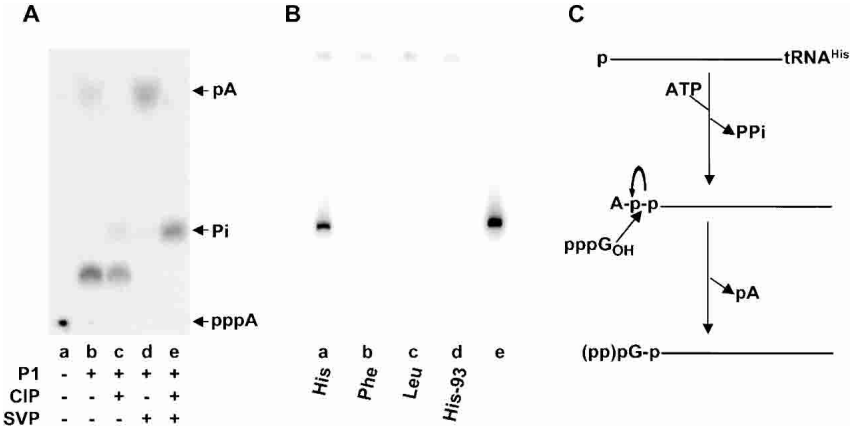

Figure 6. Analysis of reaction of His_6-Thg1p with p-tRNA and ATP. (*A*) End analysis of reaction. p-tRNA^{His} substrate was incubated with [α-^{32}P]ATP and His_6-Thg1p, as described in Materials and Methods; labeled product tRNA was gel purified, then analyzed on thin-layer plates after digestion. *a*, no enzyme; *b*, nuclease P1; *c*, nuclease P1 followed by CIP; *d*, nuclease P1 followed by snake venom pyrophosphatase (SVP); *e*, nuclease P1 followed by CIP and SVP treatment. (*B*) Adenylylation specificity of His_6-Thg1p. p-tRNA was prepared from ppp-tRNA by dephosphorylation and rephosphorylation, as described in Materials and Methods, and then assayed for adenylylation activity. *a*, p-tRNA^{His} (75 nt); *b*, p-tRNA^{Phe}; *c*, p-tRNA^{Leu}; *d*, p-tRNA^{His} (93 nt); *e*, labeled tRNA^{His} standard. (*C*) Proposed mechanism for tRNA^{His} guanylylation.

lane c); treatment with pyrophosphatase yields material that comigrates with AMP (Fig. 6A, lane d), and treatment with pyrophosphatase and phosphatase yields Pi (Fig. 6A, lane e). Thus the product of the reaction of p-tRNA with ATP and His_6-Thg1p is adenylylated tRNA.

Two lines of evidence suggest that this adenylylation is due to His_6-Thg1p. First, the ATP-dependent guanylyltransferase observed with p-tRNA^{His} copurifies with His_6-Thg1p through an additional gel filtration column (data not shown). Second, as shown in Figure 6B, the adenylylation reaction is specific for tRNA^{His} (lane a) relative to other tRNAs prepared in parallel, including p-tRNA^{Phe} (lane b), p-tRNA^{Leu} (lane c), and a 93-mer p-tRNA^{His} (His-93) substrate containing an 18-nt extended 5′ end (lane d). Thus it seems likely that the adenylylation is caused by Thg1p and not by a contaminant.

The most reasonable conclusion is that the guanylyltransferase reaction with p-tRNA^{His} is activated with ATP by adenylylation of the 5′ end, followed by the guanylyltransferase step to displace AMP (Fig. 6C). This proposed mechanism is similar to that previously reported by others for a 58-kD yeast protein, which is apparently unrelated to either Thg1p or YDR076c based on their different sizes (Jahn and Pande 1991). Indeed, we have shown that treatment of the adenylylated tRNA with GTP discharges material that comigrates with AMP (not shown). A presumed pyrophosphatase activity would yield the final tRNA^{His} product with a pG end; this is likely to follow guanylyltransfer, to account for the fact that the final product is mostly (but not completely) pG.

Thg1p lacks known enzymatic motifs yet is widely conserved within the domains Eucarya and Archaea

Although Thg1p is clearly required for tRNA^{His} guanylyltransferase activity in vivo in yeast (Fig. 3), and has demonstrable activity when purified from *E. coli* and is therefore free of other yeast proteins (Figs. 4–6), analysis of its primary structure yields no obvious functional clues that are consistent with the demonstrated activities of the protein. The sequence of Thg1p displays no known structural motifs such as those implicated in RNA/DNA binding, nucleotidyltransferase, or pyrophosphatase activity; in particular, Thg1p lacks the six motifs that are well conserved in various combinations among members of the ATP-dependent ligase family and cellular and viral capping enzymes (Shuman and Schwer 1995). A Pfam (Bateman et al. 2000) search indicates that Thg1p is a member of the DUF549 family of unknown function, whereas a CDART (Conserved Domain Architecture Retrieval Tool) search at NCBI shows that these and several other family members are the sole constituents of COG4021.

Despite the absence of obvious functional domains in Thg1p, homologs of the corresponding gene are widely distributed within eukaryotes and to a somewhat more limited degree among archaeal species, but are not identifiable in eubacteria (Supplementary Tables 1, 2). With few exceptions (notably *Caenorhabditis*), presumptive Thg1p gene orthologs (Fig. 7) are found in all free-living eukaryotes whose genomes have been completely sequenced. Within the clearly alignable portion of the sequence (yeast residues 1–219), Thg1p is on average 49% identical and 67% similar to other eukaryotic homologs, with uniformly distributed blocks of invariant (or nearly invariant) residues (Fig. 7); conservation of these blocks is decidedly lower among the archaeal representatives (see below), especially within the C-terminal half of the protein.

In Archaea, Thg1p homologs are confined almost exclusively to the Euryarchaeota. Thg1p-like genes are found in Methanosarcinales and Methanopyrales, but are not detected in the sequenced genomes of euryarchaeote species such as *Archaeoglobus fulgidus*, *Methanocaldococcus jannaschii*, or *Halobacterium* sp. With the exception of a particularly divergent sequence in *Pyrobaculum*

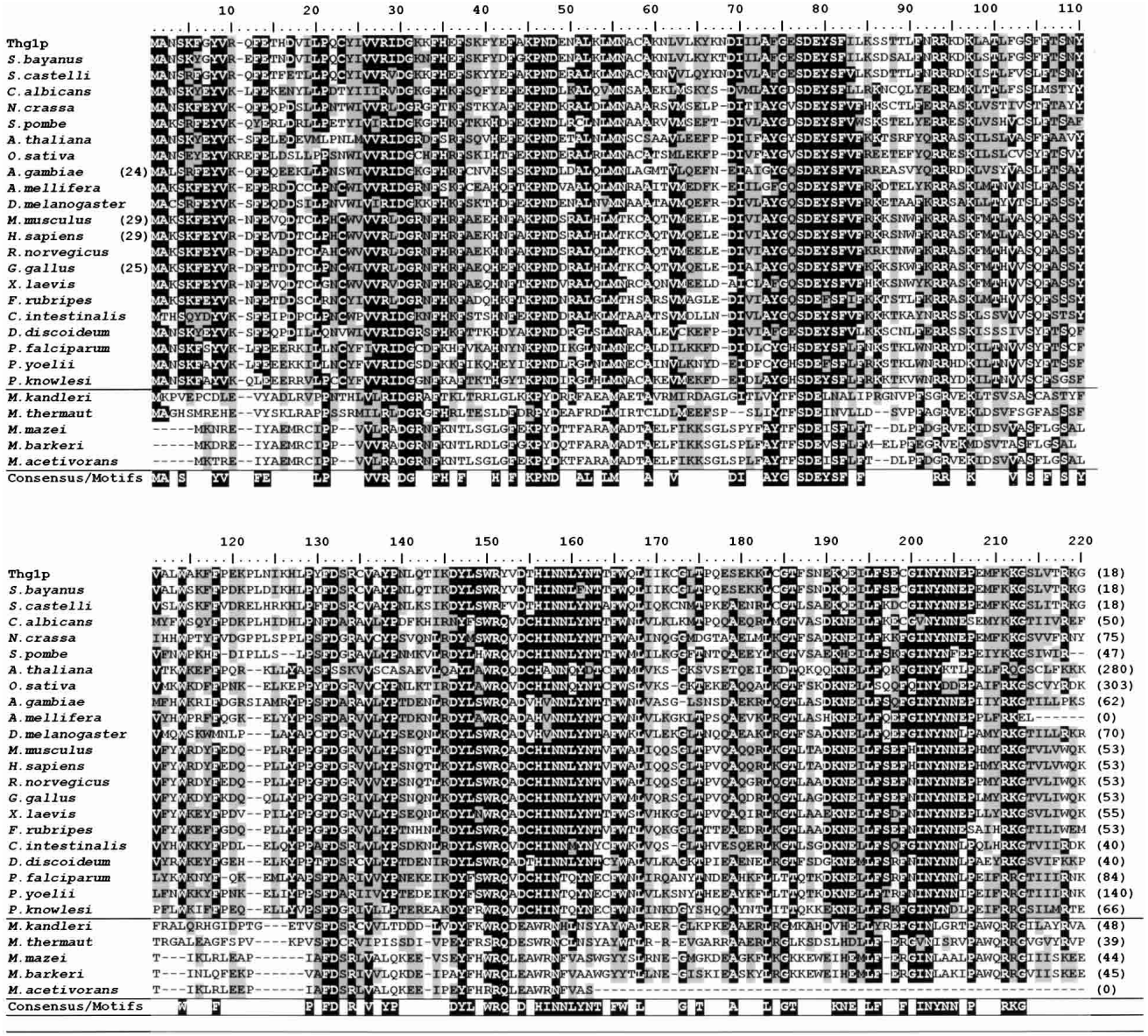

Figure 7. CLUSTALX alignment of full-length Thg1p homologs. Accession numbers and related information are compiled in Supplementary Tables 1 and 2. The archaeal Thg1p homologs are set between horizontal lines. In a few cases, putative N-terminal mitochondrial targeting peptides were identified; numbers in parentheses preceding individual sequences indicate the length of these sequences. For clarity the variable C-terminal portion of the alignment has been omitted; numbers in parentheses at end of each sequence refer to the number of residues thus excluded. The rice and *Arabidopsis* genes represent continuous ORFs, each encoding a tandem duplication (to be described elsewhere) of the Thg1p homolog, with no identifiable internal methionine for initiation of the second copy. Only the 5′ complete copy of each of the plant genes is included in the alignment. Consensus motifs are shown at the bottom of the alignment. Using BioEdit, shading is based on a 60% consensus, with identical residues depicted on a black background and similar residues on a gray background.

aerophilum, a Thg1p homolog is not detectable in sequenced genomes from the second archaeal kingdom, Crenarchaeota, and is not present in any sequenced eubacterial genome.

Discussion

We have shown that yeast ORF YGR024c, encoding Thg1p, is responsible in vivo and in vitro for the unusual guanylyltransferase activity in which an extra guanine nucleotide residue is posttranscriptionally added to the 5′ end of $tRNA^{His}$. We identified Thg1p in a survey of a genomic collection of purified GST-ORF fusion proteins for GTP incorporation activity with $tRNA^{His}$, and then proved that the product tRNA has an extra guanine nucleotide at its 5′ terminus in the correct phosphodiester linkage, primarily with a monophosphate end. We concluded that Thg1p is required in vivo for guanylyla-

tion of tRNA^{His} because a mutant strain that conditionally lacks Thg1p also conditionally lacks the extra nucleotide at the 5′ end of tRNA^{His}. Our observations further suggest that addition of this G_{-1} residue to tRNA^{His} is its essential role, based on the coincident loss of cell growth, tRNA^{His} 5′ end maturation, and guanylyltransferase activity in extracts, as cells are depleted of Thg1p.

Recombinant His_6-Thg1p, purified after expression in *E. coli*, has several activities expected for tRNA^{His} guanylyltransferase. With a ppp-tRNA^{His} substrate and GTP, this protein can catalyze the G_{-1} addition step and most or all of the removal of the pyrophosphate that occurs during formation of mature tRNA^{His}. This preparation of His_6-Thg1p also displays tRNA^{His} guanylyltransferase activity with a p-tRNA^{His} substrate, the more natural substrate expected after removal of the 5′ leader of the precursor tRNA by RNase P; this reaction requires ATP in addition to GTP, and our evidence indicates that the activation occurs through adenylylation. This adenylylation activity is likely due to Thg1p because it copurifies through a further purification step, and because it is specific for tRNA^{His}. However, we note that the guanylyltransferase reaction with p-tRNA and ATP does not reach the levels achieved with a ppp-tRNA^{His} substrate under conditions we have tested. Thus, we cannot rule out the possibility that the comparatively weak adenylylation activity is not the physiologically relevant activation step. Thg1p appears to be a novel enzyme, because there is no hint in its sequence of domains for any of the expected reactions, such as the KX(D/N)G motif implicated in the predominant family of guanylyltransferases and ligases (Shuman and Schwer 1995; Ho and Shuman 2002), or similarity to known pyrophosphatases.

There may be other components of yeast tRNA^{His} guanylyltransferase. At present we observe less than a single turnover of Thg1p with a ppp-tRNA^{His} substrate, assuming it is all active, and the ATP-dependent reaction with p-tRNA^{His} substrate is somewhat weaker. One or more of these activities could be stimulated by the protein encoded by ORF YDL076c, which also copurifies with tRNA^{His} guanylyltransferase activity (see above). However, unlike Thg1p, the protein encoded by ORF YDL076c is not essential, and its role, if any, in the reaction catalyzed by Thg1p is unknown. Another protein that may be involved is the previously identified 58-kD polypeptide that was present in a purified preparation of a tRNA^{His} guanylyltransferase activity (Pande et al. 1991). Perhaps this 58-kD protein acts in concert with Thg1p in vivo in this or another closely related role. Alternatively, this protein may be an independent tRNA^{His} guanylyltransferase activity that acts in another capacity in the cell.

Yeast tRNA^{His} guanylyltransferase is almost unique, because it catalyzes addition of a nucleotide in a normal phosphodiester linkage at the 5′ end of an RNA, formally in a 3′ to 5′ direction. Other nucleotide addition reactions almost always add to the 3′ end of an RNA, either in a templated reaction, as in polymerases, mRNA editing (Ernst et al. 2003), and likely tRNA editing reactions (Lavrov et al. 2000; Hopper and Phizicky 2003), or in a nontemplated reaction, as in mRNA poly(A) formation (Lingner et al. 1991) and tRNA CCA addition (Shi et al. 1998). The only other documented example in which phosphodiester bond formation occurs in a 3′–5′ direction is the highly unusual activity in *Acanthamoeba castellanii* that edits a number of mitochondrial tRNAs by removal of up to three unpaired nucleotides at the 5′ end, followed by templated polymerization in the 3′–5′ direction (Lonergan and Gray 1993; Price and Gray 1999). Like Thg1p, the *Acanthamoeba* activity requires activation by a molecule such as ATP when acting upon substrates bearing a 5′-phosphate end, and it exhibits no ATP dependence when presented with a 5′-triphosphorylated substrate (Price and Gray 1999). Thg1p differs somewhat from the *A. castellanii* activity because it generates a product that bears mainly a 5′-monophosphate rather than a 5′-triphosphate terminus, and because addition is likely nontemplated: G_{-1} in yeast tRNA^{His} is not expected to pair with A_{73} at the end of the acceptor stem of tRNA^{His}. Indeed, the lack of templating or the presence of A_{73} may be important for the tRNA^{His} guanylyltransferase reaction, because the widespread distribution of Thg1p homologs in eukaryotes correlates perfectly with the presence of A_{73} in their tRNA^{His} species. In contrast, the absence of Thg1p homologs in eubacteria is consistent with the presence of an encoded G_{-1} residue opposite C_{73} (Sprinzl et al. 1998; Marck and Grosjean 2002).

Of 20 yeast tRNA modification proteins that have been identified (Hopper and Phizicky 2003; Jackman et al. 2003), Thg1p is only the third one that is essential under standard laboratory conditions. Gcd10p/Gcd14p catalyzes formation of m^1A_{58} in a number of tRNAs (Anderson et al. 1998, 2000), and is essential for modification of $tRNA_i^{Met}$ (Calvo et al. 1999), and Tad2p/Tad3p catalyzes formation of I_{34} in the wobble position of the anticodon of tRNA^{Ala} (Gerber and Keller 1999), and is presumably essential for proper decoding during translation. Strains lacking each of three other modification enzymes are viable, but have distinct growth defects (Lecointe et al. 1998; Bjork et al. 2001; Pintard et al. 2002), whereas lack of most other modification proteins has only subtle growth defects.

Thg1p may be essential because its absence leads to a lack of charged His-tRNA^{His} in the cell. The G_{-1} residue is a crucial determinant for yeast histidyl-tRNA synthetase. Absence of G_{-1} lowers k_{cat}/K_m of histidyl-tRNA synthetase by ~500–700-fold, whereas alterations at anticodon positions 34 and 35 in tRNA^{His} are only responsible for a 20–70-fold effect on k_{cat}/K_m (Rudinger et al. 1994; Nameki et al. 1995). Furthermore, presence of G_{-1} in the context of a different tRNA raises the k_{cat}/K_m for histidyl-tRNA synthetase by 80–230-fold (Rudinger et al. 1994; Nameki et al. 1995). Thus, absence of G_{-1} may lead to a lethal level of undercharging of tRNA^{His} in the cell. We note that the corresponding G_{-1}:C_{73} base pair of bacterial tRNA^{His} is also an important determinant of synthetase recognition, although the essential contributions

to tRNA identity may be from C_{73} and the phosphate at position −1 (Francklyn and Schimmel 1990; Yan and Francklyn 1994; Fromant et al. 2000).

Why do all sequenced $tRNA^{His}$ species have an extra G_{-1} residue, and how did the activity that adds G_{-1} originate? One could imagine that the inclusion of G_{-1} arose originally from the pairing to C_{73} in eubacteria, and the anomalous cleavage at position −1 by RNase P to leave the extra base pair in the acceptor stem (Orellana et al. 1986; Burkard et al. 1988). Subsequent steps could involve the evolution of a variant $tRNA^{His}$ species that had an A_{73} residue opposite G_{-1}, which might have been cleaved less efficiently at G_{-1} by RNase P (Orellana et al. 1986; Burkard et al. 1988) but selectively favored during translation; and the recruitment of an activity that could catalyze guanylyl transfer for those $tRNA^{His}$ molecules that were cleaved at the +1 position. In this regard, a further intriguing question is whether there is any direct evolutionary relationship (i.e., evidence of homology) between the functionally similar activities that carry out $tRNA^{His}$ guanylylation in yeast (i.e., Thg1p) and mitochondrial 5′-tRNA editing in *A. castellanii*. Identification of the gene(s) encoding the *A. castellanii* activity should shed light on this possibility. Similarly, elucidation of the catalytic mechanism of Thg1p may illuminate the origin of this unusual activity.

Materials and methods

Preparation of substrate tRNAs

To make tRNA substrates beginning with 5′ triphosphate (ppp-tRNA), plasmids were linearized by digestion with NsiI or Bst NI, and transcribed at 37°C for 90 min in reaction mixtures containing 50 mM Tris-HCl (pH 8.0), 30 mM $MgCl_2$, 1 mM spermidine, 7.5 mM DTT, 2.0 mM each NTP, 0.1 mg/mL template DNA, and 20–50 µg/mL T7 RNA polymerase. Transcripts were purified from a 10% polyacrylamide gel containing 4 M urea.

Three methods were used to make tRNA substrates beginning with 5′ monophosphate (p-tRNA). In most cases p-tRNA was generated by inclusion of 20 mM GMP during in vitro transcription, which results in transcripts mostly beginning with p-tRNA (Himeno et al. 1989; Sampson et al. 1989). For the experiment shown in Figure 5B, substrate beginning only with p-tRNA was prepared by transcription of pGu14 to produce tRNA with 18 additional nucleotides (GAACCCTGTGCAAG CAAC) at its 5′ end, followed by gel purification of the product, and treatment with RNase H in the presence of a chimeric 2′-*O*-methylated/deoxy oligonucleotide (HHMI/Keck Center, Yale Univ.) to direct cleavage 3′ of the extra nucleotides at the 5′ end (Jackman et al. 2003; Yu 1999). The p-$tRNA^{His}$ product was then gel-purified. For the experiment shown in Figure 6B, all p-tRNA substrates were prepared from in vitro-transcribed ppp-tRNA, by phosphatase treatment and subsequent kinase treatment. Approximately 420 pmole ppp-tRNA was incubated in 50-µL reaction mixtures containing 0.1 U/µL calf intestinal phosphatase in buffer (Roche) at 55°C for 1 h, followed by heat inactivation at 95°C for 3 min, two phenol/chloroform extractions, and ethanol precipitation. Dephosphorylation efficiency was calculated by addition of trace amounts of [5′-^{32}P]-labeled $tRNA^{His}$ in parallel reactions. Dephosphorylated tRNA was then rephosphorylated in 60-µL reaction mixtures containing 1 U/µL polynucleotide kinase in buffer (Roche), 5 mM ATP, at 37°C for 45 min. Rephosphorylation efficiency was calculated by inclusion of 4 µM [γ-^{32}P]ATP (final specific activity 5.6 Ci/mmol) in parallel reactions. Then RNA was extracted with phenol/chloroform, precipitated twice, resolved on a 10% polyacrylamide gel containing 4 M urea, and purified.

Assay for $tRNA^{His}$ guanylyltransferase activity

Activity was assayed based on the incorporation of [α-^{32}P]GTP into tRNA (Cooley et al. 1982) in 10–40-µL reaction mixtures containing 20–25 mM HEPES (pH 7.5–7.9), 10 mM $MgCl_2$, 100–125 mM KCl or NaCl, 3 mM DTT, 0.4 µM tRNA, 0.4–0.8 µM [α-^{32}P]GTP (400 Ci/mmole), 0–3 mM ATP, and protein. Reactions were incubated at room temperature for 90 min, and RNA was extracted with phenol/chloroform, precipitated with ethanol, and resolved on a 10% polyacrylamide gel containing 4 M urea.

Adenylylation of p-tRNA

Adenylylation of p-tRNA was assayed based on the incorporation of [α-^{32}P]ATP into p-tRNA in 10-µL reaction mixtures containing 25 mM HEPES (pH 7.5), 10 mM $MgCl_2$, 125 mM NaCl, 3 mM DTT, 0.4–0.8 µM p-tRNA, 2.5–3.0 µM [α-^{32}P]ATP (400 Ci/mmol), 10 µM ATP, and 2–4 µM Thg1p. Reactions were incubated at room temperature for 60 min, and RNA was extracted with phenol/chloroform, precipitated twice, resolved on a 10% polyacrylamide gel containing 4 M urea, and purified when stated.

Preparation of yeast extracts and pools of purified GST-ORF fusion proteins from yeast

Yeast crude extracts were prepared using glass beads in buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 4 mM $MgCl_2$, 5 mM DTT, 10% glycerol, 1 M NaCl, and protease inhibitors (McCraith and Phizicky 1990; Phizicky et al. 2002). Samples were quick-frozen and stored at −80°C. Yeast S100 extracts were prepared essentially as described (Klekamp and Weil 1982). Pools and subpools of purified GST-ORF fusion proteins were prepared as described (Martzen et al. 1999; Phizicky et al. 2002).

Plasmids

The $tRNA^{His}$ gene (comprising nt +1 to the CCA end of mature tRNA) was PCR-amplified from yeast DNA with forward primer (5′AAAACTGCAGTAATACGACTCACTATAGCCA TCTTAGTATAG) containing a PstI site, a T7 promoter, and tRNA sequence starting from G_{+1}, and with reverse primer (5′CCCGCGGATCCATGCATGGTGCCATCTCCTAGAATC) containing a BamHI site, an NsiI site (for runoff transcription), and tRNA sequence starting from the CCA end. Amplified DNA was inserted into a pUC13 vector to generate plasmid pGu3. Plasmid pGu2 contains mature $tRNA^{His}$ sequence (G_{-1} to the CCA end), and plasmid pGu14 contains the sequence GAACCCTGTGCAAGCAAC followed by G_{+1} to the CCA end; both plasmids were constructed in the same way as pGu3. Plasmids containing mature $tRNA^{Phe}$ or mature $tRNA^{Leu}$ sequence were constructed by RT-PCR of mature tRNA spliced in vitro, with a BstN1 site downstream for runoff transcription. Plasmid pGu11 was constructed to purify Thg1p (YGR024c), by insertion of genome-amplified *THG1* sequence into a pET-derived vector (pBG1861) specifying the N-terminal sequence Met-Ala-His_6 followed by the initiation codon (gift from E. Grayhack,

University of Rochester Medical School). Plasmid pGu4 (*CEN URA3* P_{GAL10}-*THG1*) was constructed by inserting the *THG1* gene into vector pAVA0040 (Alexandrov et al. 2002) between BamHI and PstI sites.

End analysis of gel-purified RNA

3′ end-labeling of tRNA^{His} tRNA^{His} was labeled with [5′-^{32}P]pCp at its 3′ end using T4 RNA ligase as described (Butler et al. 1997).

P1 nuclease RNA was incubated in reaction mixtures containing 10 mM NaAc (pH 5.3), 0.2 µg/µL P1 (Calbiochem), 2 µg/µL bulk RNA, and 0.2 mM $ZnCl_2$ at 37°C for 30 min. Products were resolved by thin-layer chromatography on polyethyleneimine (PEI) cellulose (EM Science) using 0.5 M $NaHCO_2$ (pH 3.4) as solvent.

Snake venom pyrophosphatase (SVP) treatment RNA was incubated in reaction mixtures containing 10 mM Tris-Cl (pH 9.0), 14 mM $MgCl_2$, 0.05 U/µL SVP (Sigma), 1 mM NAD at 37°C for 30 min, followed by phenol/choloroform extraction, and products were resolved by thin-layer chromatography on PEI cellulose using 0.5 M $NaHCO_2$ (pH 3.4) as solvent.

Calf intestinal phosphatase (CIP) treatment RNA was incubated in 5-µL reaction mixtures containing dephosphorylation buffer (Roche), 2 µg/µL yeast bulk RNA, and 0.1 U/µL CIP (Roche) at 37°C for 30 min, and products were resolved by thin-layer chromatography on PEI cellulose using 1 M LiCl as solvent (Fig. 2B) or 0.5 M $NaHCO_2$ (Fig. 6A). In Figure 6A, SVP buffer was used for RNA codigestion with CIP and SVP.

Partial RNase T1 digestion RNA was incubated in buffer containing 16 mM sodium acetate (pH 5.0), 5.6 M urea, 0.42 µg/µL bulk RNA, and 0.008 U/µL RNase T1 (Industrial Research) at 50°C for 20 min, and products were resolved on a 12% polyacrylamide gel containing 7 M urea (Kuchino and Nishimura 1989).

Growth of E. coli *and purification of His_6-Thg1p*

Cells were grown in 1 l LB containing ampicillin at 37°C to A^{600} = 0.6, induced at 18°C overnight by the addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG), and cells were harvested and frozen at −70°C. His_6-Thg1p was purified by immobilized metal-ion affinity chromatography (IMAC) as described (Jackman et al. 2003), using TALON resin (Clontech).

Yeast strains

Deletion strains and wild-type haploid parents were purchased from Invitrogen: strain 26977 (*MATa/α, his3–Δ1/ his3–Δ1, leu2–Δ0/ leu2–Δ0, met15–Δ0/ met15–Δ0, ura3–Δ0/ ura3–Δ0, YGR024/ygr024–Δ0*); BY4742 (*MATα, his3–Δ1, leu2–Δ0, met15–Δ0, ura3–Δ0*).

To construct a strain conditionally lacking Thg1p, yeast strain BY4742 was transformed with plasmid pGu4 (*CEN URA3* P_{GAL10}-*YGR024*) to obtain strain WG12. Then a DNA fragment containing the *thg1*-Δ::kan^r cassette and flanking sequences (from −271 to +273 relative to the ORF ends) was PCR-amplified from genomic DNA of strain 26977 and transformed into strain WG12 to generate strain WG18 (relevant genotype: *ygr024-Δ0*::kan^rP_{GAL10}-*YGR024*), after selection on media containing galactose and 200 µg/mL G418.

Preparation of small-molecular-weight RNA

RNA was extracted from 2×10^9 cells with hot phenol, precipitated twice with ethanol, and resuspended in 250 µL 10 mM Tris-HCl (pH 7.5) as described (Rubin 1975). The concentration of purified RNA was calculated by assuming 1 unit of A^{260} = 40 µg/mL RNA.

Primer extension analysis

Primers corresponding to the dihydrouridine loop region (5′ GA TGTGTACTAACCACTAT for tRNA^{His}, and 5′ TAAAAGC CGAACGCTCTACC for tRNA$^{Lys}_{UUU}$) were used for primer extension after 5′ end-labeling with T4 polynucleotide kinase and [γ-^{32}P]ATP (ICN, 7000 Ci/mmol), and purification, as reported (Jackman et al. 2003). Annealed primer was extended for 5 min at room temperature, followed by 1 h at 37°C with 20 µM dNTP and 0.4 U/µL AMV-reverse transcriptase in AMV-RT reaction buffer (Promega). Sequencing reactions additionally contained 10 µM of each corresponding ddNTP. Reactions were stopped by addition of equal volume of RNA loading buffer (80% formamide, 1 mM EDTA). Products were resolved by 12% PAGE with 4 M urea, and visualized by PhosphorImager (Molecular Dynamics).

Acknowledgments

We thank Beth Grayhack, Neil Shull, Andrei Alexandrov, and Feng Xing for helpful advice throughout this work and during preparation of this manuscript. This research was supported by NIH grants GM52347 and HG02311 to E.M.P., by grant MOP-4124 to M.W.G. from the Canadian Institutes of Health Research, and by NIH postdoctoral fellowship 2T32CA09363 to J.E.J.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Alexandrov, A.V., Martzen, M.R., and Phizicky, E.M. 2002. Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. *RNA* **8:** 1253–1266.

Anderson, J., Phan, L., Cuesta, R., Carlson, B.A., Pak, M., Asano, K., Bjork, G.R., Tamame, M., and Hinnebusch, A.G. 1998. The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes & Dev.* **12:** 3650–3662.

Anderson, J., Phan, L., and Hinnebusch, A.G. 2000. The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **97:** 5173–5178.

Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., and Sonnhammer, E.L. 2000. The Pfam protein families database. *Nucleic Acids Res.* **28:** 263–266.

Bjork, G.R., Jacobsson, K., Nilsson, K., Johansson, M.J., Bystrom, A.S., and Persson, O.P. 2001. A primordial tRNA modification required for the evolution of life? *EMBO J.* **20:** 231–239.

Bullerwell, C.E., Forget, L., and Lang, B.F. 2003. Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res.* **31:** 1614–1623.

Burkard, U., Willis, I., and Soll, D. 1988. Processing of histidine transfer RNA precursors. Abnormal cleavage site for RNase P. *J. Biol. Chem.* **263:** 2447–2451.

Butler, J.S., Briggs, M.W., and Proweller, A. 1997. Analysis of polyadenylation phenotypes in *Saccharomyces cerevisiae*. In *mRNA formation and function* (ed. J.D. Richter), pp. 111–125. Academic Press, New York.

Calvo, O., Cuesta, R., Anderson, J., Gutierrez, N., Garcia-Barrio, M.T., Hinnebusch, A.G., and Tamame, M. 1999. GCD14p, a repressor of GCN4 translation, cooperates with Gcd10p and Lhp1p in the maturation of initiator methionyl-tRNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19:** 4167–4181.

Cooley, L., Appel, B., and Soll, D. 1982. Posttranscriptional nucleotide addition is responsible for the formation of the 5′ terminus of histidine tRNA. *Proc. Natl. Acad. Sci.* **79:** 6475–6479.

Ernst, N.L., Panicucci, B., Igo Jr., R.P., Panigrahi, A.K., Salavati, R., and Stuart, K. 2003. TbMP57 is a 3′ terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol. Cell* **11:** 1525–1536.

Francklyn, C. and Schimmel, P. 1990. Enzymatic aminoacylation of an eight-base-pair microhelix with histidine. *Proc. Natl. Acad. Sci.* **87:** 8655–8659.

Fromant, M., Plateau, P., and Blanquet, S. 2000. Function of the extra 5′-phosphate carried by histidine tRNA. *Biochemistry* **39:** 4062–4067.

Gerber, A.P. and Keller, W. 1999. An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **286:** 1146–1149.

Himeno, H., Hasegawa, T., Ueda, T., Watanabe, K., Miura, K., and Shimizu, M. 1989. Role of the extra G-C pair at the end of the acceptor stem of tRNA(His) in aminoacylation. *Nucleic Acids Res.* **17:** 7855–7863.

Ho, C.K. and Shuman, S. 2002. Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains. *Proc. Natl. Acad. Sci.* **99:** 12709–12714.

Hopper, A.K. and Phizicky, E.M. 2003. tRNA transfers to the limelight. *Genes & Dev.* **17:** 162–180.

Jackman, J.E., Montange, R.K., Malik, H.S., and Phizicky, E.M. 2003. Identification of the yeast gene encoding the tRNA m^1G methyltransferase responsible for modification at position 9. *RNA* **9:** 574–585.

Jahn, D. and Pande, S. 1991. Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae*. II. Catalytic mechanism. *J. Biol. Chem.* **266:** 22832–22836.

Klekamp, M.S. and Weil, P.A. 1982. Specific transcription of homologous class III genes in yeast-soluble cell-free extracts. *J. Biol. Chem.* **257:** 8432–8441.

Kuchino, Y. and Nishimura, S. 1989. Enzymatic RNA sequencing. *Methods Enzymol.* **180:** 154–163.

L'Abbe, D., Lang, B.F., Desjardins, P., and Morais, R. 1990. Histidine tRNA from chicken mitochondria has an uncoded 5′-terminal guanylate residue. *J. Biol. Chem.* **265:** 2988–2992.

Laforest, M.-J., Roewer, I., and Lang, B.F. 1997. Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG "stop" codons recognized as leucine. *Nucleic Acids Res.* **25:** 626–632.

Lavrov, D.V., Brown, W.M., and Boore, J.L. 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl. Acad. Sci.* **97:** 13738–13742.

Lecointe, F., Simos, G., Sauer, A., Hurt, E.C., Motorin, Y., and Grosjean, H. 1998. Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of Ψ_{38} and Ψ_{39} in tRNA anticodon loop. *J. Biol. Chem.* **273:** 1316–1323.

Lingner, J., Kellermann, J., and Keller, W. 1991. Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. *Nature* **354:** 496–498.

Lonergan, K.M. and Gray, M.W. 1993. Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. *Science* **259:** 812–816.

Marck, C. and Grosjean, H. 2002. tRNomics: Analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* **8:** 1189–1232.

Martzen, M.R., McCraith, S.M., Spinelli, S.L., Torres, F.M., Fields, S., Grayhack, E.J., and Phizicky, E.M. 1999. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286:** 1153–1155.

McCraith, S.M. and Phizicky, E.M. 1990. A highly specific phosphatase from *Saccharomyces cerevisiae* implicated in tRNA splicing. *Mol. Cell. Biol.* **10:** 1049–1055.

Nameki, N., Asahara, H., Shimizu, M., Okada, N., and Himeno, H. 1995. Identity elements of *Saccharomyces cerevisiae* tRNA(His). *Nucleic Acids Res.* **23:** 389–394.

Orellana, O., Cooley, L., and Soll, D. 1986. The additional guanylate at the 5′ terminus of *Escherichia coli* $tRNA^{His}$ is the result of unusual processing by RNase P. *Mol. Cell. Biol.* **6:** 525–529.

Pande, S., Jahn, D., and Soll, D. 1991. Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae*. I. Purification and physical properties. *J. Biol. Chem.* **266:** 22826–22831.

Phizicky, E.M., Consaul, S.A., Nehrke, K.W., and Abelson, J. 1992. Yeast tRNA ligase mutants are nonviable and accumulate tRNA splicing intermediates. *J. Biol. Chem.* **267:** 4577–4582.

Phizicky, E.M., Martzen, M.R., McCraith, S.M., Spinelli, S.L., Xing, F., Shull, N.P., Van Slyke, C., Montagne, R.K., Torres, F.M., Fields, S., et al. 2002. Biochemical genomics approach to map activities to genes. *Methods Enzymol.* **350:** 546–559.

Pintard, L., Lecointe, F., Bujnicki, J.M., Bonnerot, C., Grosjean, H., and Lapeyre, B. 2002. Trm7p catalyses the formation of two 2′-O-methylriboses in yeast tRNA anticodon loop. *EMBO J.* **21:** 1811–1820.

Price, D.H. and Gray, M.W. 1999. A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in *Acanthamoeba castellanii*. *RNA* **5:** 302–317.

Riley, L.G., Ralston, G.B., and Weiss, A.S. 1996. Multimer formation as a consequence of separate homodimerization domains: The human c-Jun leucine zipper is a transplantable dimerization module. *Protein Eng.* **9:** 223–230.

Rubin, G.M. 1975. Preparation of RNA and ribosomes from yeast. *Methods Cell Biol.* **12:** 45–64.

Rudinger, J., Florentz, C., and Giege, R. 1994. Histidylation by yeast HisRS of tRNA or tRNA-like structure relies on residues -1 and 73 but is dependent on the RNA context. *Nucleic Acids Res.* **22:** 5031–5037.

Sampson, J.R., DiRenzo, A.B., Behlen, L.S., and Uhlenbeck, O.C. 1989. Nucleotides in yeast $tRNA^{Phe}$ required for the specific recognition by its cognate synthetase. *Science* **243:** 1363–1366.

Schnare, M.N., Heinonen, T.Y., Young, P.G., and Gray, M.W. 1985. Phenylalanine and tyrosine transfer RNAs encoded by *Tetrahymena pyriformis* mitochondrial DNA: Primary sequence, post-transcriptional modifications, and gene localization. *Curr. Genet.* **9:** 389–393.

Shi, P.Y., Maizels, N., and Weiner, A.M. 1998. CCA addition by tRNA nucleotidyltransferase: Polymerization without translocation? *EMBO J.* **17:** 3197–3206.

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

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. 1998. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26:** 148–153.

Williams, J.B., Cooley, L., and Soll, D. 1990. Enzymatic addition of guanylate to histidine transfer RNA. *Methods Enzymol.* **181:** 451–462.

Xing, F., Martzen, M.R., and Phizicky, E.M. 2002. A conserved family of *Saccharomyces cerevisiae* synthases effects dihydrouridine modification of tRNA. *RNA* **8:** 370–381.

Yan, W. and Francklyn, C. 1994. Cytosine 73 is a discriminator nucleotide in vivo for histidyl-tRNA in *Escherichia coli*. *J. Biol. Chem.* **269:** 10022–10027.

Yu, Y.T. 1999. Construction of 4-thiouridine site-specifically substituted RNAs for cross-linking studies. *Methods* **18:** 13–21.

$tRNA^{His}$ maturation: An essential yeast protein catalyzes addition of a guanine nucleotide to the 5′ end of $tRNA^{His}$

Weifeng Gu, Jane E. Jackman, Amanda J. Lohan, et al.

Genes Dev. 2003, **17:**
Access the most recent version at doi:10.1101/gad.1148603