

# tRNA<sup>His</sup>-guanylyltransferase establishes tRNA<sup>His</sup> identity

Ilka U. Heinemann<sup>1</sup>, Akiyoshi Nakamura<sup>2</sup>, Patrick O'Donoghue<sup>1</sup>, Daniel Eiler<sup>1</sup> and Dieter Söll<sup>1,3,\*</sup>

<sup>1</sup>Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, USA,

<sup>2</sup>Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan and <sup>3</sup>Department of Chemistry, Yale University, New Haven, CT 06520-8114, USA

Received July 12, 2011; Revised August 9, 2011; Accepted August 10, 2011

## ABSTRACT

Histidine transfer RNA (tRNA) is unique among tRNA species as it carries an additional nucleotide at its 5' terminus. This unusual G<sup>-1</sup> residue is the major tRNA<sup>His</sup> identity element, and essential for recognition by the cognate histidyl-tRNA synthetase to allow efficient His-tRNA<sup>His</sup> formation. In many organisms G<sup>-1</sup> is added post-transcriptionally as part of the tRNA maturation process. tRNA<sup>His</sup> guanylyltransferase (Thg1) specifically adds the guanylate residue by recognizing the tRNA<sup>His</sup> anticodon. Thg1 homologs from all three domains of life have been the subject of exciting research that gave rise to a detailed biochemical, structural and phylogenetic enzyme characterization. Thg1 homologs are phylogenetically classified into eukaryal- and archaeal-type enzymes differing characteristically in their cofactor requirements and specificity. Yeast Thg1 displays a unique but limited ability to add 2–3 G or C residues to mutant tRNA substrates, thus catalyzing a 3' → 5' RNA polymerization. Archaeal-type Thg1, which has been horizontally transferred to certain bacteria and few eukarya, displays a more relaxed substrate range and may play additional roles in tRNA editing and repair. The crystal structure of human Thg1 revealed a fascinating structural similarity to 5' → 3' polymerases, indicating that Thg1 derives from classical polymerases and evolved to assume its specific function in tRNA<sup>His</sup> processing.

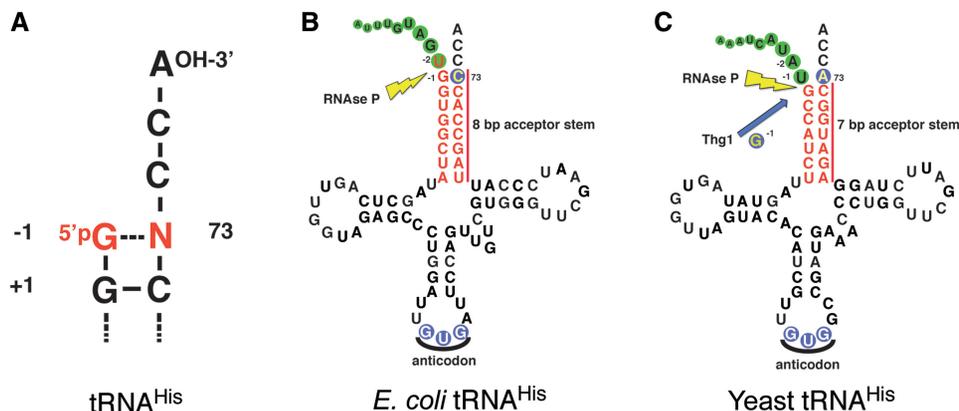
## INTRODUCTION

The astonishing process of translation, which converts an informational RNA string composed of only four different nucleotide bases into functional proteins containing up

to 22 different cotranslationally inserted amino acids, relies on the flawless decoding of the genetic information. During translation, tRNAs serve as adapter molecules, deciphering genetic code triplets from messenger RNA into amino acids. The fidelity of this process requires the correct ligation of a particular amino acid on its cognate tRNA. Several strategies are employed by the aminoacyl tRNA synthetases (aaRS) to ensure that only their cognate tRNAs are subjected to aminoacylation. While all tRNA species exhibit a highly conserved secondary and L-shaped tertiary structure, subtle variations in the sequence and structure of tRNA molecules allow the aaRSs to differentiate cognate substrates from non-substrates. The specific nucleotides and structural motifs that are recognized by the cognate aaRS are called identity elements. Common identity elements are the anticodon, the discriminator base, modified bases in the anticodon loop and varying bases in the loop and stem structures. While some identity elements serve as determinants, signaling “correct substrate” to the aaRS, others prevent misaminoacylation by acting as anti-determinants (1).

The correct aminoacylation of tRNA<sup>His</sup>, in almost all organisms, relies on the presence of a unique additional guanylate base at the 5'-end of tRNA<sup>His</sup>, the so called G<sup>-1</sup> residue (2–9). In bacteria, G<sup>-1</sup> basepairing with the cytidine discriminator base extends the regular 7 base pair acceptor stem into an 8-base pair stem (Figure 1), which is otherwise only found in tRNA<sup>Sec</sup>. Histidyl-tRNA synthetases (HisRS) from all three kingdoms recognize this specific G<sup>-1</sup> residue and depend on its presence for efficient aminoacylation. In the absence of G<sup>-1</sup>, tRNA<sup>His</sup> histidylation is reduced to 1% *in vitro*, which is not sufficient for survival *in vivo* (2,4,5). A tRNA microhelix consisting of the acceptor stem including G<sup>-1</sup> only, has been shown to be sufficient to promote aminoacylation by *Escherichia coli* HisRS *in vitro* (7). Thus, G<sup>-1</sup> presence in tRNA<sup>His</sup> is the primary identity element for most HisRS and vital to retain translational fidelity.

\*To whom correspondence should be addressed. Tel: +1 203 432 6200; Fax: +1 203 432 6202; Email: dieter.soll@yale.edu



**Figure 1.** tRNA<sup>His</sup> contains an unusual G<sup>-1</sup> residue. (A) The monophosphate of the tRNA<sup>His</sup>ΔG<sup>-1</sup> residue is the major identity element of HisRS and highly conserved throughout the three kingdoms of life. The discriminator base 73 can either be a cytidine (in archaea and bacteria) or adenosine (in eukarya) residue. (B) and (C) RNase P displays two alternate cleavage patterns for tRNA<sup>His</sup>: (B) In bacteria, the G<sup>-1</sup> residue forms an additional base pair with a C<sup>73</sup> discriminator base, extending the acceptor stem to 8 base pairs. RNase P cleaves at the -1 residue, yielding mature tRNA<sup>His</sup> containing a 5'-monophosphate. (C) In eukarya, RNase P cleaves at its regular +1 cleavage site, generating tRNA<sup>His</sup> lacking G<sup>-1</sup> with a 5'-monophosphate at the G<sup>+1</sup> residue. The G<sup>-1</sup>-residue is not genome-encoded, but post-transcriptionally added by Thg1.

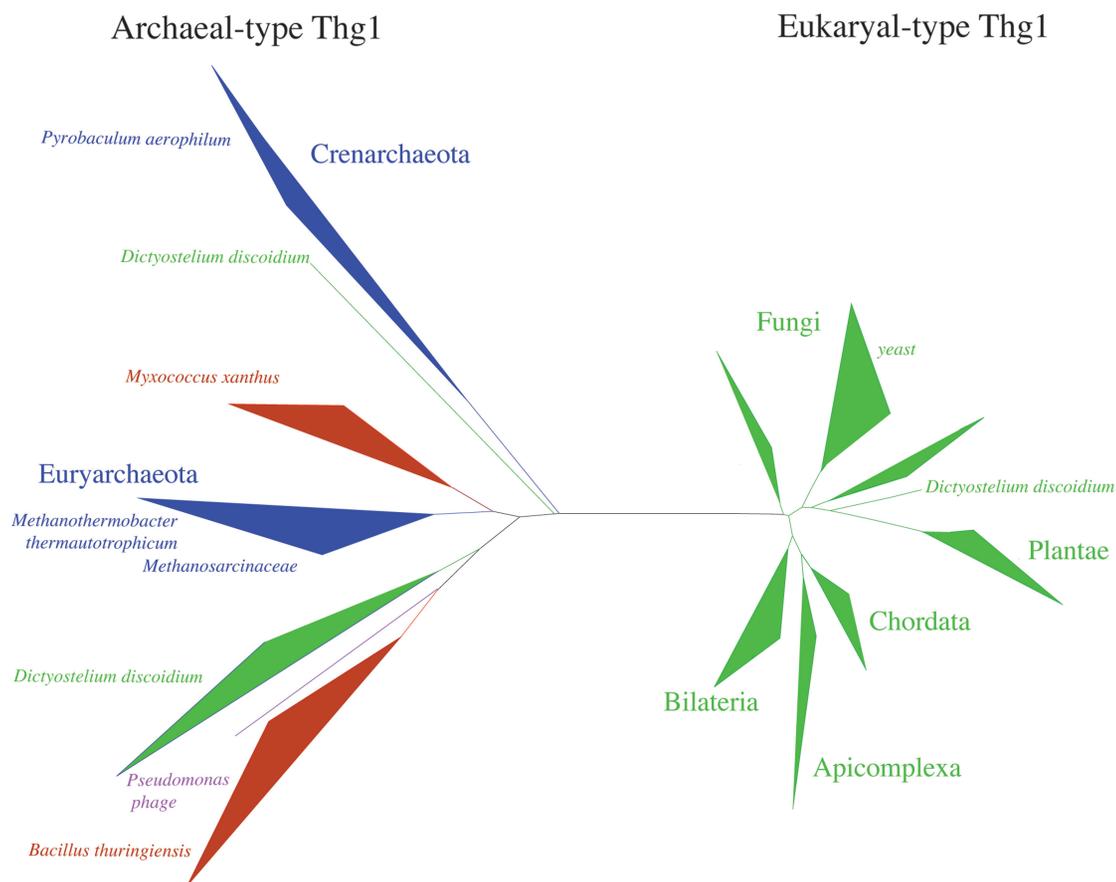
Intriguingly, two separate pathways for providing tRNA<sup>His</sup> with the crucial G<sup>-1</sup> residue are found in nature (Figure 1). Bacteria, except a small  $\alpha$ -proteobacterial clade discussed below, generally encode the G<sup>-1</sup> residue in their genome and rely on an aberrant RNase P cleavage to produce active tRNA<sup>His</sup>. While RNase P usually reliably cleaves at the +1 position of tRNAs, the enzyme displays an altered cleavage pattern for tRNA<sup>His</sup> (10,11). This unusual cleavage pattern is thought to be promoted by the primary acceptor stem structure, with the additional G<sup>-1</sup> and C<sup>73</sup> base paired and a uridine at position -2 immediately upstream of the RNase P cleavage site being of primary importance ((12) and references therein) (Figure 1). Thus, RNase P cleaves the precursors for all tRNAs, including tRNA<sup>His</sup>, correctly.

Thus it came as a surprise to find histidine tRNAs from a small  $\alpha$ -proteobacterial clade, e.g. *Caulobacter crescentus*, completely lack the G<sup>-1</sup> residue (13,14). However, the same organisms also contain a HisRS with a number of small peptide insertions. Anticodon recognition is possibly due to an insertion of an anticodon recognition loop in the amino acid sequence of these HisRSs (13–15). In a clear case of co-evolution between tRNA and synthetase, the unusual HisRS is only found in the  $\alpha$ -proteobacteria that dispensed with the G<sup>-1</sup> residue on tRNA<sup>His</sup>.

In contrast, eukarya employ a different way to provide tRNA<sup>His</sup> with its identity element. In the eukaryal genomes sequenced so far, the G<sup>-1</sup> residue is not encoded in the genome, and an adenosine in position 73 opposite G<sup>-1</sup> does not allow for base pairing. Thus, RNase P cleaves in its regular fashion at the +1 residue. Nevertheless, eukaryal HisRSs rely on the presence of G<sup>-1</sup> in mature tRNA to promote aminoacylation (4). The guanine residue is therefore enzymatically added in a 3'→5' polymerization reaction after tRNA processing by RNase P. While the presence of the G<sup>-1</sup> residue in tRNA<sup>His</sup> in

*Drosophila melanogaster* and *Schizosaccharomyces pombe* has been demonstrated almost 30 years ago (16), and the enzymatic activity was purified later (17–19), the open reading frame encoding the tRNA<sup>His</sup> guanylyltransferase (Thg1) activity was only identified recently (20). Since this identification, vast progress has been made in the function and role of Thg1 through protein characterization from all three domains of life.

While Bacteria and Eukarya generally seem to adhere to their method of ensuring the presence of the tRNA<sup>His</sup> identity element, the situation in archaea is less clear. Most archaea genomically encode the G<sup>-1</sup> residue and are thus likely to retain G<sup>-1</sup> after RNase P cleavage (21). Other archaea likely rely on the post-transcriptional G<sup>-1</sup> addition by Thg1, since G<sup>-1</sup> is not genome encoded. In these cases a Thg1 homolog is present in the genome (Figure 2). Interestingly, some archaea, and also few bacteria, encode both G<sup>-1</sup> and Thg1, making it less clear in which way G<sup>-1</sup> is provided. The same applies to some mitochondrial and chloroplast tRNA<sup>His</sup> species, which potentially encode a G<sup>-1</sup>, but in which a post-transcriptional G<sup>-1</sup> addition by a chromosome-encoded Thg1 with mitochondrial targeting sequence is possible (22–24). In plant mitochondria, the coexistence of both pathways has been experimentally demonstrated *in vivo*. *Arabidopsis thaliana* mitochondrial RNase P cleaves pre-tRNA<sup>His</sup> containing G<sup>-1</sup> in two positions, creating a mixture of mature tRNA<sup>His</sup> and tRNA<sup>His</sup> lacking G<sup>-1</sup> (thereafter termed tRNA<sup>His</sup>ΔG<sup>-1</sup>). The latter is then guanylated to yield tRNA<sup>His</sup>. So far, the identity of the respective guanylyltransferase is unclear, since both *A. thaliana* Thg1 homologs localize in the cytoplasm (25). Whether an unrelated enzyme fulfills the guanylation reaction, or tRNA<sup>His</sup>ΔG<sup>-1</sup> is exported into the cytoplasm for further processing, is unclear. The data suggest that even when G<sup>-1</sup> is genome encoded, Thg1 is retained to repair miscleaved tRNA<sup>His</sup> produced by RNase P, indicating why Thg1 has been selectively retained during evolution.



**Figure 2.** Phylogeny of Thg1: the figure was adapted from (29). Organisms are color-coded. Red, Bacteria; Blue, Archaea; Green, Eukarya; Purple, Viruses. Archaeal and eukaryal-type Thg1 show a deep phylogenetic divide. The archaeal-type Thg1 consists of a highly diverse group with homologs from the archaea, bacteria, eukarya and viruses, whereas eukaryal-type Thg1 is exclusively found in eukarya.

## THG1 ACTIVITY

### The crucial $G^{-1}$ of tRNA<sup>His</sup> in the three domains of life

**Eukaryal Thg1.** The primary Thg1 function is undoubtedly the addition of a guanylate residue to tRNA<sup>His</sup> $\Delta G^{-1}$  to provide tRNA<sup>His</sup> with its primary identity element. Initially, the open reading frame encoding Thg1 activity was identified by characterization of a yeast *thg1* knockout, which accumulates uncharged tRNA<sup>His</sup> $\Delta G^{-1}$  and displays a growth defect (20,26). While a diploid *thg1* deletion strain containing only one gene copy is still viable, a complete Thg1 deletion is lethal in yeast (20). The lack of Thg1 in a yeast *thg1* deletion mutant can to some extent be compensated by overexpression of HisRS and tRNA<sup>His</sup> $\Delta G^{-1}$ , making the residual HisRS charging activity on tRNA<sup>His</sup> $\Delta G^{-1}$  sufficient for survival (27). Yeast Thg1 has, since its discovery, been studied in great detail, leading to the proposal of a reaction mechanism involving an adenosine triphosphate (ATP) cofactor for primary activation, anticodon and discriminator base recognition, and a final guanylation step to yield mature tRNA<sup>His</sup> (summarized in Table 1).

The addition of the crucial guanylate residue by Thg1 is highly specific for tRNA<sup>His</sup> with the major identity

element being the anticodon and the discriminator base (28–30). Interestingly, HisRS, unlike most aaRS, does not recognize the anticodon and discriminator base of mature tRNA<sup>His</sup>. Instead, HisRS identifies tRNA<sup>His</sup> chiefly by the presence of the  $G^{-1}$  residue. It remains unclear why many organisms employ two enzymes (HisRS and Thg1) and not simply a single enzyme, as in the anticodon-recognizing HisRS from *C. crescentus*, to accurately form His-tRNA<sup>His</sup>. The situation leads to speculation regarding an important regulatory function of pre-tRNA<sup>His</sup> and Thg1. Different studies have demonstrated, that uncharged tRNA can be used to regulate gene expression under stress conditions (31–33), but no such function has been assigned to tRNA<sup>His</sup> $\Delta G^{-1}$  so far.

**Archaeal Thg1.** The first characterized Thg1 homolog from archaea based on its activity *in vivo* and *in vitro* was from *Methanosarcina acetivorans* (21). While *M. acetivorans* and other sequenced Methanosarcina species genomically encode the  $G^{-1}$  residue in tRNA<sup>His</sup> gene, they additionally contain an open reading frame for Thg1. *M. acetivorans* cell extracts display Thg1 activity and the pure enzyme is active *in vitro* (21), but it is unclear whether *M. acetivorans* relies on Thg1 for  $G^{-1}$

**Table 1.** Eukaryal-type and archaeal-type Thg1 homologs differ in substrate requirements and reaction specificity

	Eukaryal-type Thg1	Archaeal-type Thg1
Substrate requirements		
Nucleotide used for tRNA activation	ATP	ATP or GTP
Anticodon recognition	Yes, for p-tRNA substrates	Yes, for p-tRNA substrates
Discriminator base ( <sup>73</sup> ) of natural tRNA substrate	Adenine	Cytidine
Complementation of yeast <i>thg1</i> knockout	Yes	Yes, usually a C <sup>73</sup> -tRNA <sup>His</sup> substrate is required
Discriminator base-dependent NTP addition		
Discriminator base A <sup>73</sup>	Physiological substrate. (GTP), (ATP), (UTP), (CTP). Only GTP is added in competition experiments	UTP, inefficient GTP addition
Discriminator base C <sup>73</sup>	GTP, ATP, UTP, CTP	Physiological substrate. GTP
Discriminator base G <sup>73</sup>	GTP, ATP, UTP, CTP	CTP
Discriminator base U <sup>73</sup>	GTP, ATP, UTP, CTP	ATP
Further reactions with mutated tRNA <sup>His</sup> substrates		
Reverse polymerization beyond G <sup>-1</sup>	Requires base pairing, limited to GTP and CTP	Observed for Hyperthermophiles, limited to multiple GTP addition
N <sup>+1</sup> addition	Little or none	Templated, all NTPs
N <sup>+2</sup> addition	Not determined	Templated, GTP and CTP
Substrate tRNAs other than tRNA <sup>His</sup>	No	Different tRNAs for N <sup>+1</sup> addition (tRNA <sup>Ile</sup> and tRNA <sup>Leu</sup> )

addition, or whether RNase P displays its unusual cleavage pattern. Whether Thg1 activity in *M. acetivorans* is essential or Thg1 serves as a backup mechanism as suggested for plant mitochondria (25), remains to be elucidated. Other archaeal Thg1 homologs are more likely to be essential to retain tRNA<sup>His</sup> fidelity. Enzymatic activities for *Methanothermobacter thermautotrophicus*, *Methanopyrus kandleri* and *Pyrobaculum aerophilum* Thg1 have been experimentally verified *in vitro* (29,34). None of these species encode a G<sup>-1</sup> residue on the chromosome (21). While the sequence identity between archaeal and eukaryal enzymes is rather low (archaeal and eukaryal Thg1s consist of 15% identical residues on average), enzymes from both kingdoms add the G<sup>-1</sup> to tRNA<sup>His</sup> (20,21,29,34). Even though the Thg1 general activity is the same in both archaea and eukarya, some mechanistic differences involving cofactor dependence and tRNA recognition elements have been observed.

**Bacterial Thg1.** Only recently, the biochemical Thg1 characterization has been extended to homologs from bacteria (29,30). As mentioned above, bacteria do not seem to require Thg1. While a few bacteria encode an unusual HisRS capable of anticodon recognition, the vast majority of bacteria possess a genome encoded G<sup>-1</sup> residue that is retained following RNase P cleavage, yet several bacteria from different clades encode Thg1 homologs (29). Although apparently not required for active tRNA<sup>His</sup> production, tRNA<sup>His</sup> guanylyltransferase activity of *Bacillus thuringiensis* and *Myxococcus xanthus* Thg1 homologs has been demonstrated (29). These results were confirmed in an independent study shortly thereafter (30). While several bacterial homologs have been studied *in vitro*, their *in vivo* function has not been subject to investigation. It is

currently unknown, whether the bacterial homologs are essential to the cell or entirely dispensable, and whether their *in vivo* function is restricted to tRNA<sup>His</sup>ΔG<sup>-1</sup> guanylation.

**Thg1 Classification.** A detailed phylogenetic analysis of Thg1 homologs from all domains of life revealed that while most eukaryal and archaeal Thg1 homologs cluster according to accepted phylogeny, bacterial Thg1 homologs are interspersed in the archaeal clade (Figure 2) (29). Within this clade two major archaeal groups are found, separated into crenarchaeota and euryarchaeota, whereas bacterial homologs are distributed rather randomly. Thg1 occurrence in bacteria is therefore most likely the result of at least two independent horizontal gene transfers from archaea to bacteria and further horizontal gene transfer among bacterial lineages (29). The phylogeny suggests a general Thg1 classification into eukaryal and archaeal Thg1 types (Figure 2). Biochemical data add further support to this classification. For example, eukaryal-type Thg1 is strictly ATP dependent and requires an A<sup>73</sup> discriminator base for G<sup>-1</sup> addition, while archaeal-type Thg1 can utilize either ATP or guanosine triphosphate (GTP) for activation and requires a C<sup>73</sup> discriminator base (Table 1).

## SUBSTRATE REQUIREMENTS

The primary Thg1 substrates are the respective tRNA<sup>His</sup>ΔG<sup>-1</sup> and GTP to yield mature tRNA<sup>His</sup> and inorganic pyrophosphate (PP<sub>i</sub>). The yeast enzyme requires ATP to activate the tRNA substrate prior to G<sup>-1</sup> addition (20). Thus, Thg1 has to differentiate tRNA<sup>His</sup> from non-cognate tRNA species, substrate tRNA<sup>His</sup>ΔG<sup>-1</sup>

from product tRNA<sup>His</sup>, recognize ATP and distinguish GTP from other nucleotides in order to promote specific tRNA<sup>His</sup>ΔG<sup>-1</sup> guanylation. The two major identity elements for Thg1 are apparently the anticodon and the discriminator base of tRNA<sup>His</sup>ΔG<sup>-1</sup>. Thg1 most likely acts on tRNA substrates with a fully processed 3'-end, as it displays activity on respective *in vitro* transcribed tRNAs with a 3'CCA-end. A participation of the 3'-end of tRNA<sup>His</sup>ΔG<sup>-1</sup> in the reaction has however not been demonstrated.

### tRNA<sup>His</sup> identity elements—anticodon recognition

Thg1 needs to specifically recognize and edit tRNA<sup>His</sup>ΔG<sup>-1</sup> and provide it with the guanylate residue, thus producing a suitable substrate for HisRS. An accidental guanylation of a different tRNA would also make other tRNAs containing the G<sup>-1</sup> sufficient substrates for HisRS, as shown for a tRNA<sup>Ala</sup> microhelix (4,35). A shared Thg1 feature throughout the three domains of life is the tRNA<sup>His</sup> anticodon recognition. While tRNA<sup>Phe</sup> is a poor substrate for Thg1, a transplantation of the His-anticodon GUG to the tRNA<sup>Phe</sup> body promotes guanylation by Thg1 (28,29). Also, replacing the tRNA<sup>His</sup> anticodon with the Phe anticodon GAA drastically decreases Thg1 enzymatic activity, substantiating the anticodon as major identity element for Thg1. For yeast Thg1, the anticodon discrimination plays a role in the recognition of its natural substrate monophosphate tRNA<sup>His</sup> (p-tRNA<sup>His</sup>ΔG<sup>-1</sup>), which is the result of proper RNase P cleavage. Results obtained in Thg1 assays with *in vitro* transcribed tRNA substrates containing an unnatural 5'-triphosphate moiety differ significantly in the discrimination of the anticodon. In the absence of guanidine monophosphate (GMP), *in vitro* transcribed tRNAs have a triphosphate at the 5'-end of tRNA<sup>His</sup> (ppp-tRNA<sup>His</sup>ΔG<sup>-1</sup>). The tRNA<sup>His</sup> substrate terminology is outlined in Figure 3 and its figure legend. While p-tRNA<sup>Phe</sup> is not a substrate for yeast Thg1, ppp-tRNA<sup>Phe</sup> is readily guanylated by Thg1 *in vitro* (28). ppp-tRNA substrates apparently mimic the activated tRNA<sup>His</sup>ΔG<sup>-1</sup> substrates and do not allow the enzyme to distinguish the Phe-anticodon from the His-anticodon. Thus, the enzyme only differentiates the anticodon when provided with the physiological substrate p-tRNA, indicating that both the anticodon and the 5'-monophosphate are important for yeast Thg1 to distinguish substrate from non-substrate tRNA, but at different stages of the reaction.

### Templated versus non-templated G<sup>-1</sup> addition—the discriminator base

While anticodon recognition is a shared feature among the various Thg1 homologs, eukaryal- and archaeal-type Thg1 largely differ in their nucleotide preference for the discriminator base. According to accepted tRNA base numbering, the discriminator base is referred to as base number 73 (36). In a non-templated reaction, yeast Thg1 adds G<sup>-1</sup> opposite to A<sup>73</sup>, which is conserved in all eukarya (20). But, provided with only one nucleotide at a time, and under high enzyme concentrations, yeast Thg1 can add any nucleotide to tRNA<sup>His</sup> with A<sup>73</sup> (37). When the yeast tRNA<sup>His</sup> discriminator base is altered to C<sup>73</sup>, U<sup>73</sup>

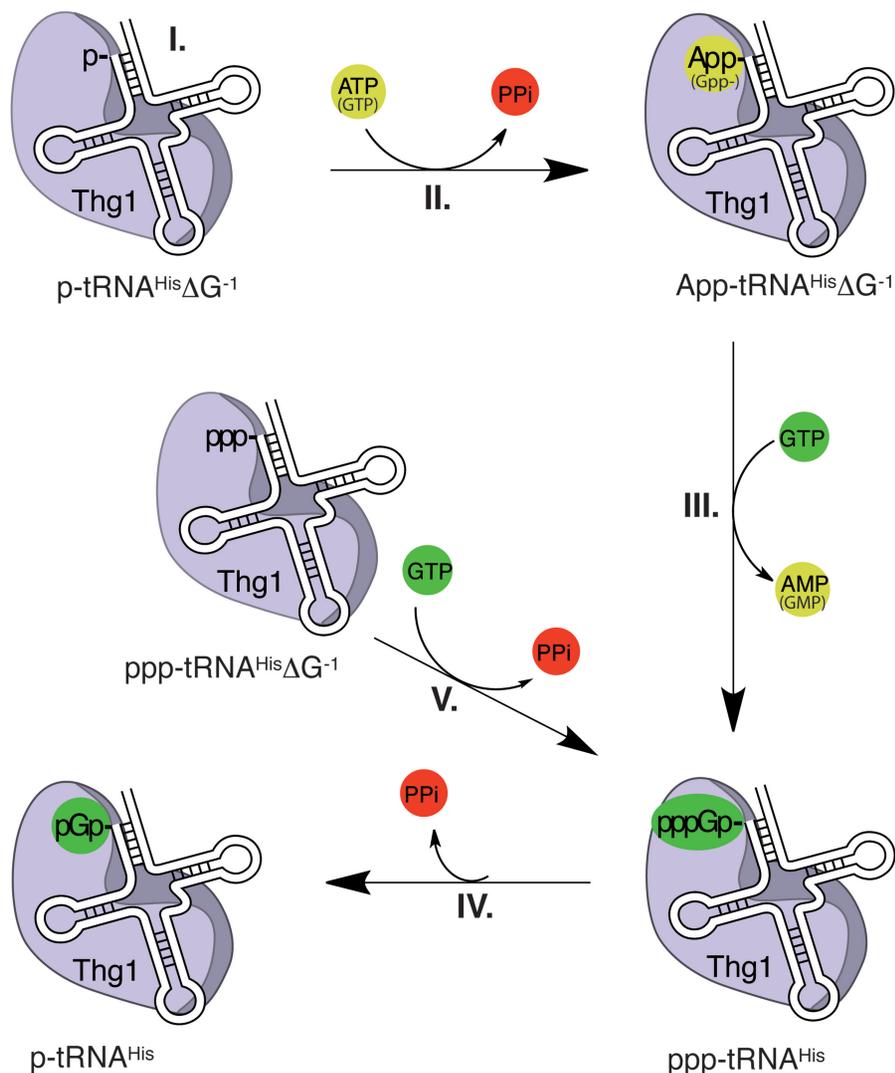
or G<sup>73</sup>, Thg1 in high concentrations can add any nucleotide at the -1 position, as shown by experiments using mutated tRNA<sup>His</sup> substrates (37). The mutation of the yeast tRNA<sup>His</sup> discriminator base to C<sup>73</sup> furthermore leads to a prolonged extension reaction; producing a tRNA<sup>His</sup> with up to three added guanine residues. The presence of an adenine at position 73 is thus highly important to restrict Thg1 activity to the addition of a single guanine residue, only. In the presence of all four nucleotides and wild-type tRNA<sup>His</sup>, Thg1 exclusively adds GTP to wt-tRNA<sup>His</sup>ΔG<sup>-1</sup> (37). Thus, the wild-type reaction catalyzed by yeast Thg1 is the non-templated GTP addition.

Archaea, in contrast, require a cytidine residue as a discriminator base and support efficient guanylation of their homologous tRNA substrates (21). In a non-homologous system, archaeal Thg1 could add G<sup>-1</sup> to mutant yeast tRNA with C<sup>73</sup>, but not to wild-type yeast tRNA<sup>His</sup> containing A<sup>73</sup> (34). This also explains the inability of archaeal Thg1 homologs to complement a yeast *thg1* deletion mutant *in vivo*, unless a suitable tRNA<sup>His</sup>ΔG<sup>-1</sup> substrate containing C<sup>73</sup> is provided (34). Yeast Thg1 activates the tRNA substrate with ATP, which has been shown to be a prerequisite for guanylation (Figure 3). Archaeal homologs from *M. kandleri*, *M. acetivorans* and *Methanosarcina barkeri* Thg1 efficiently activate both yeast tRNA<sup>His</sup> with A<sup>73</sup> or a mutant yeast tRNA<sup>His</sup> with C<sup>73</sup>. This suggests that the activation step is independent of the discriminator base recognition (34).

Some bacterial Thg1 homologs appear to be less stringent in discriminator base recognition. In most bacteria tRNA<sup>His</sup> encode C<sup>73</sup> that may aid the unusual RNase P cleavage, which leaves the G<sup>-1</sup> position intact (Figure 1). While *in vitro* the *B. thuringiensis* homolog appears to prefer a C<sup>73</sup> for efficient guanylation, it does promote residual guanylation on an A<sup>73</sup> containing yeast tRNA<sup>His</sup> substrate (30). *In vivo* complementation experiments revealed that *B. thuringiensis* Thg1 complements a yeast *thg1* knockout and restores growth to near wild-type levels (29). The data show, that *B. thuringiensis* Thg1 accepts A<sup>73</sup>-tRNA<sup>His</sup> as substrate *in vivo*. These results were confirmed in an independent study that demonstrated a relaxed discriminator base specificity of bacterial Thg1s (30). All Thg1 activities observed to date are summarized in Table 1.

### Cofactor requirements

The guanylation reaction as carried out by yeast Thg1 requires the tRNA<sup>His</sup>ΔG<sup>-1</sup> activation by ATP. An adenylated tRNA<sup>His</sup>ΔG<sup>-1</sup> intermediate has been trapped (App-tRNA<sup>His</sup>ΔG<sup>-1</sup>), suggesting a primary activation step upon ATP hydrolysis and a secondary displacement of the adenylate by guanylate to yield AMP and mature tRNA<sup>His</sup> (37) (Figure 3). Employment of ppp-tRNA<sup>His</sup>ΔG<sup>-1</sup> as substrate, which mimics the adenylated intermediate, may circumvent the requirement of the initial adenylation reaction (see Figure 3 for a description of the terminology used for tRNA substrates and intermediates). Interestingly, the substrate discrimination from non-substrate tRNAs seems to occur at or before the initial adenylation step. The anticodon recognition



**Figure 3.** Reaction mechanism of Thg1: tRNA<sup>HisΔG<sup>-1</sup></sup>, tRNA<sup>His</sup> lacking the G<sup>-1</sup> residue; p-tRNA, 5'-monophosphorylated tRNA; ppp-tRNA, 5'-triphosphorylated tRNA; App-tRNA and Gpp-tRNA, 5'-activated adenylated or guanylated tRNA. (I) Thg1 binds monophosphorylated tRNA<sup>HisΔG<sup>-1</sup></sup> lacking the G<sup>-1</sup> residue (p-tRNA<sup>HisΔG<sup>-1</sup></sup>). (II) p-tRNA<sup>HisΔG<sup>-1</sup></sup> can be activated with either ATP or GTP to yield App-tRNA<sup>HisΔG<sup>-1</sup></sup> or Gpp-tRNA<sup>HisΔG<sup>-1</sup></sup>. (III) Activated substrate tRNA is guanylated with GTP to yield triphosphorylated tRNA<sup>His</sup> (ppp-tRNA<sup>His</sup>). (IV) The final dephosphorylation step leads to mature monophosphorylated tRNA<sup>His</sup> (p-tRNA<sup>His</sup>). (V) The activation step can be bypassed with 5'-triphosphorylated substrate-tRNA (ppp-tRNA<sup>His</sup>).

is less stringent for ppp-tRNA (600-fold preference of ppp-tRNA<sup>HisΔG<sup>-1</sup></sup> over ppp-tRNA<sup>Phe</sup>) substrates compared to p-tRNA substrates (10 000-fold preference of p-tRNA<sup>HisΔG<sup>-1</sup></sup> over p-tRNA<sup>Phe</sup>) (28). While there is no known physiological ppp-tRNAs source in the cell, this result suggests that anticodon recognition is carried out prior or simultaneously to the adenylation reaction. The final guanylation step then requires either the physiologically relevant 5'-adenylated tRNA or a 5'-triphosphate-tRNA substrate.

In archaeal-type Thg1, ATP is not required for the reaction and is presumably replaced by GTP (29). Archaeal-type enzymes efficiently catalyze the guanylation reaction in the absence of ATP. In the presence of ATP and GTP however, a reaction with *B. thuringiensis* Thg1 leads to App-tRNA<sup>HisΔG<sup>-1</sup></sup> accumulation, but not Gpp-tRNA<sup>HisΔG<sup>-1</sup></sup>. The result indicates, that activation

by ATP is preferred to activation by GTP (30). For archaeal-type Thg1 the initial ATP-dependent activation step may be carried out using GTP as cofactor in a reaction consuming two GTP molecules per tRNA<sup>His</sup>. This scenario is further substantiated by the recently published human Thg1 crystal structure (detailed below). In the crystal structure, the likely ATP-binding pocket is occupied by dGTP, showing that GTP can sterically fit into the yeast ATP binding pocket and may very well replace ATP in the reaction (29,38). The data to this point indicates that mechanistic differences exist between archaeal and eukaryal-type Thg1.

#### Pyrophosphatase activity

The efficient aminoacylation by HisRS requires a tRNA<sup>His</sup> with a G<sup>-1</sup> residue containing a 5'-monophosphate. In

*E. coli* HisRS, a cluster of arginine residues in the active site recognizes this monophosphate group. The major role of the  $G^{-1}$  residue in HisRS recognition is to correctly position the 5'-monophosphate in the HisRS active site (7-9). Little is known on how the 5'-triphosphate, which is generated as result of the guanylation reaction, is dephosphorylated. Presumably Thg1 possesses pyrophosphatase activity, since the major reaction product formed in the *in vitro* reaction is p-tRNA<sup>His</sup>. Yet some unprocessed ppp-tRNA<sup>His</sup> has been observed (20). The precise mechanism is unclear, but dephosphorylation is either accomplished as a final step or in concert with GTP transfer to the activated tRNA (Figure 3). In bacteria, the enzyme RppH converts ppp-tRNA substrate to p-tRNA and in this way initiates RNA degradation (39,40). Whether Thg1 utilizes a similar mechanism for PP<sub>i</sub> removal as RppH remains to be investigated.

## ACTIVITIES BEYOND $G^{-1}$ ADDITION

### Addition of other nucleotides and extension beyond $G^{-1}$

The only demonstrated and biologically significant *in vivo* Thg1 function is the tRNA<sup>His</sup>  $\Delta G^{-1}$  guanylation. In Thg1 *in vitro* assays, homologs from both yeast and archaea have shown that Thg1 can add other nucleotides than G to non-natural substrates (summarized in Table 1). Yeast Thg1 can add any nucleotide in the  $G^{-1}$  position, independent of the discriminator base. In competition experiments with all four NTPs, GTP is strongly preferred as a substrate (37). Beyond  $G^{-1}$  addition, however, the "reverse polymerization" behavior of yeast Thg1 appears to be template dependent, but is restricted to G-C base pairing. In experiments with tRNA variants, a  $G^{-2}$  residue was added to a tRNA with the typical  $C^{74}$ , while a  $C^{-2}$  was added opposite to a mutated tRNA with a  $G^{74}$ . The templated nucleotide addition beyond the  $-1$  position extends up to position  $-3$ , apparently halted by a 3' terminal adenosine residue. An A or U addition beyond the  $-1$  position has not been shown (37). This interesting reverse polymerization activity displayed on non-natural substrates might be a residual activity derived from an ancestral reverse polymerase function of a now highly evolved enzyme, which now specifically carries out a single non-templated guanylation reaction.

Archaeal Thg1 homologs are also capable of adding other nucleotides to a tRNA substrate. In contrast to the non-templated yeast Thg1 reaction,  $G^{-1}$  addition by archaeal Thg1 is carried out as a templated reaction matching  $G^{-1}$  to the  $C^{73}$  discriminator base. Archaeal Thg1 can add other nucleotides when the discriminator base is mutated. Opposite to  $A^{73}$  archaeal Thg1s add uridine triphosphate (UTP) in a reaction that is slightly less efficient than the guanylation reaction (34). Non-templated ATP addition opposite to a  $C^{73}$  discriminator base was not observed (29). While archaeal Thg1 homologs to some extent also catalyze the elongation beyond  $G^{-1}$ , the reaction is less prominent than in yeast Thg1 (34). Interestingly, *P. aerophilum* Thg1 promotes a significant extension reaction with multiple G residues on

a tRNA<sup>His</sup> substrate already containing the  $G^{-1}$  residue (29).

Like their archaeal counterparts, from which they are derived, bacterial Thg1 appear to strongly prefer a templated nucleotide addition (30). While all nucleotides can be added in a templated reaction, G and C addition is significantly more efficient than A and U addition. This situation resembles the preference of yeast Thg1 for G and C base addition beyond position  $-1$  (30).

### Editing of tRNA substrates

An interesting feature of archaeal-type Thg1 is its ability to display a tRNA repair function *in vitro*. *B. thuringiensis* and *M. thermautotrophicus* Thg1 are able to catalyze nucleotide addition to truncated tRNA<sup>His</sup> at the  $G^{+1}$  position, adding all four nucleotides in a templated reaction with almost equal efficiency (30). Even though archaeal-type Thg1 is clearly able to perform a repair of truncated tRNA<sup>His</sup> species, no obvious endogenous substrates for such a reaction have yet been found. A possible Thg1 function in the repair of damaged or misprocessed tRNA was therefore suggested (30). While this reaction has not yet been shown to occur *in vivo*, these *in vitro* results extend the Thg1 function from tRNA editing to 5' tRNA repair and open the possibility of additional functions for the archaeal-type Thg1 in general.

Most excitingly, this Thg1 repair function could be demonstrated for two Thg1 homologs from *Dictyostelium discoideum*, where a biological role is more plausible. *D. discoideum* encodes a total of four Thg1 homologs (29). While one of the homologs shares a high sequence similarity with yeast Thg1 (DdiTLP1), the other three homologs are phylogenetically closely related to archaeal Thg1 homologs (Figure 2). These Thg1 homologs in *D. discoideum* are most likely the result of separate lateral gene transfers from archaea to *D. discoideum* (29). Thus, separate functions for the Thg1 homologs have been suggested (41): While the eukaryal-type Thg1 homolog fulfills the classical  $G^{-1}$  addition to tRNA<sup>His</sup>  $\Delta G^{-1}$  in the cytoplasm, at least two of the archaeal-type proteins could participate in a 5' tRNA repair function. Some mitochondrial *D. discoideum* tRNAs contain mismatched bases in the tRNA acceptor stem and a 5' tRNA editing reaction of mismatched tRNA bases has been demonstrated in other Protista (42,43). The authors investigated Thg1 repair function on two naturally occurring mismatched acceptor stems, which have been truncated for the mismatched bases. Two archaeal-type Thg1 homologs were shown to add nucleotides to a truncated 5' tRNA end and thus might play a role in a repair function in coordination with an enzyme that specifically removes mismatched bases from the 5'-end of the tRNA acceptor stem. A templated repair function was demonstrated up to the  $+2$  position of truncated tRNA substrates, and included the cytidine and guanine incorporation into putative physiological substrates. While the *in vitro* results are convincing and indicate that tRNA repair might indeed be the function of these Thg1 homologs, it would be interesting to see if the repair function occurs *in vivo* by verifying the presence of such edited tRNAs in the cell and to

see if the repair function can be extended to adenosine and uridine residues.

The tRNA repair function of archaeal-type homologs explains Thg1 presence in organisms with genome encoded  $G^{-1}$ . Nevertheless,  $G^{-1}$  addition is still an important feature of archaeal-type Thg1, as some archaea do not genomically encode  $G^{-1}$  (29) and rely on Thg1 activity to produce mature tRNA<sup>His</sup>.

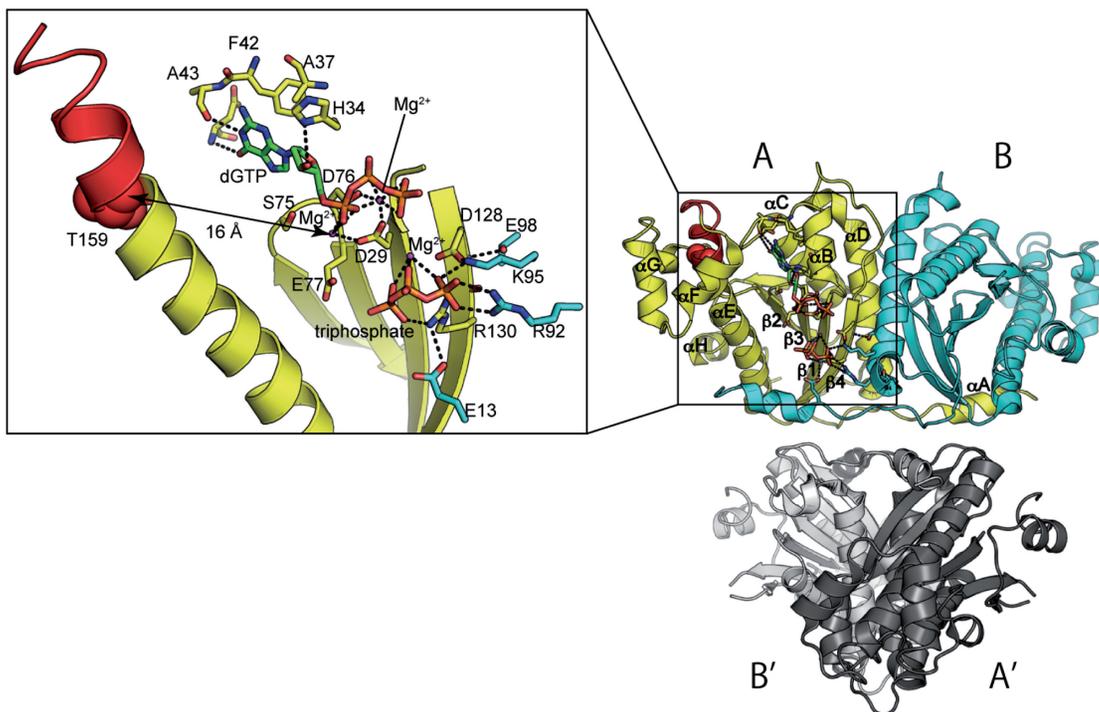
In plant mitochondria, the coexistence of both pathways has been demonstrated (25). Here, despite the ability of RNase P to leave a genome encoded  $G^{-1}$  containing tRNA<sup>His</sup> precursor intact, the presence of both tRNA<sup>His</sup> and tRNA<sup>His</sup> $\Delta G^{-1}$  have been shown as a result of two different RNase P cleavage positions. A so far unidentified Thg1 homolog then repairs the miscleaved tRNA<sup>His</sup> by adding  $G^{-1}$ . Thus, whether in organisms encoding  $G^{-1}$ , Thg1 solely acts as a repair enzyme on tRNA<sup>His</sup> $\Delta G^{-1}$ , or whether the enzyme assumes additional roles in tRNA editing remains to be elucidated.

### THE CRYSTAL STRUCTURE OF THG1

The human Thg1 crystal structure has recently been determined at 2.3 Å resolution (PDB IDs: 3OTB, 3OTC and 3OTD) (38). Thg1 shares striking structural homology with canonical 5'→3' DNA polymerases and adenylyl/guanylyl cyclases (38,44), and contains the typical highly conserved acidic residues (D, D and E) within these

enzymes families. The data indicate that Thg1 shares a similar active site and likely catalytic mechanism with other RRM-fold polymerase palm domains (45). The superposition of the three strictly conserved Thg1 carboxylates with those of T7 DNA Polymerase and their crucial role in catalysis strongly suggest that Thg1 unexpectedly uses the two-metal-ion mechanism of canonical 5'→3' polymerases (38,46).

The human Thg1 forms a tetrameric structure apparently constituted of a dimer of dimers (Figure 4). Mutational studies of yeast Thg1 suggest that alteration of strictly conserved residues in the dimer interface region strongly diminish  $G^{-1}$  addition activity (46). Thus, the tetrameric structure observed in the Thg1 crystal structure is most likely the biological unit. A tetrameric or dimeric biological unit might explain how the rather small Thg1 protein achieves its multiple tasks. Simultaneous anticodon recognition and enzymatic activity on the acceptor stem, two events happening at opposite tRNA ends, suggested a distribution of these tasks between subunits of a multimeric enzyme. In many eukaryotes, especially plants, Thg1 is encoded as a tandem protein, covalently connecting two Thg1 monomers to a dimer (29). This further underlines the likelihood of a Thg1 as a dimeric or tetrameric biological unit. In lieu of a co-crystal structure with substrate tRNA or respective biochemical data, the stoichiometry and binding mode between tetrameric Thg1 and tRNA<sup>His</sup> $\Delta G^{-1}$  remains to be elucidated.



**Figure 4.** Tetrameric structure of human Thg1 with dGTP and triphosphate (pdb 3OTB). Monomers are colored as follows: yellow, monomer A; cyan, monomer B; gray, monomer A'; white, monomer B'. The catalytic carboxylates (D29, D76 and E77), the dGTP and triphosphate recognizing residues, and the dimer interacting residues are shown as stick models. A bound dGTP and triphosphate are also shown as stick models. Three magenta spheres indicate three  $Mg^{2+}$  ions interacting with dGTP or triphosphate. The entire peptide containing pyrrolysine (amino acids O142-G153 in *M. acetivorans* corresponding to amino acids 159–169 in the human enzyme) is colored by red. The C $\alpha$  atom of T159 corresponding O142 in *M. acetivorans* (red space-filling model) is 16 Å away from  $Mg^{2+}$  ion coordinating with the catalytic carboxylates.

## Substrate binding

In the human Thg1 structure two nucleotide-binding pockets occupied by dGTP have been identified, showing the binding site of the activating nucleotide and the transferred nucleotide (Figure 4) (38). Residues interacting with the nucleotides are displayed in Figure 4. The presence of dGTP in what is most likely the ATP-binding pocket agrees with biochemical data showing that GTP can replace ATP in the activation step (29). In the crystal structure, the base and sugar moieties of the second GTP are not seen in the electron density map, most likely because of the lack of specific interactions (38). Several strictly conserved residues coordinate the observable triphosphate moiety and, if mutated to alanine, reduce catalytic activity (46). The low electron density of the sugar and base moiety indicate that for correct positioning GTP may either interact with a substrate tRNA<sup>His</sup>ΔG<sup>-1</sup> or that conformational changes induced by binding of tRNA<sup>His</sup>ΔG<sup>-1</sup> to Thg1 induce tighter GTP coordination.

tRNA binding to Thg1 is not apparent from the crystal structure, since no obvious positively charged binding groove is present. Several highly conserved residues critical to Thg1 activity are located in a small α-helical subdomain consisting of α-helices G and F (Figure 4). These subdomains appear not to play a role in catalytic activity but might be involved in tRNA binding (38). Another α-helix termed αE, lies in close proximity to the potential tRNA-binding domain, but is most likely not essential for substrate binding. This region is homologous to a pyrrolysine containing peptide in *M. acetivorans* Thg1, which has been shown to be entirely dispensable for catalytic activity (21) and therefore is most likely not involved in tRNA binding. The region homologous to the “Pyl-peptide” of *M. acetivorans* Thg1 is highlighted in red in Figure 4. Furthermore, a single aspartate residue (D68) is thought to be involved in anticodon recognition (46). This residue is located at the protein surface rather far from the active site. Anticodon recognition by this residue would either require a drastic conformational change to allow catalytic activity to be carried out by the same monomer or require dimer cooperation. How Thg1 binds tRNA<sup>His</sup> and recognizes the anticodon cannot be deduced from the structural data available. Additional biochemical investigation along with a co-crystal structure of Thg1 and its substrate will resolve this question.

## EVOLUTION OF THG1

The G<sup>-1</sup> residue is nearly universally conserved as a major identity element for recognition by HisRS. As noted, there are two pathways to produce tRNA<sup>His</sup> with G<sup>-1</sup>: (i) the aberrant RNase P cleavage that leaves a genome encoded G<sup>-1</sup> intact, (ii) the post-transcriptional G<sup>-1</sup> addition to tRNA<sup>His</sup>ΔG<sup>-1</sup> (Figure 1). It is difficult a priori to understand why two pathways exist to ensure tRNA<sup>His</sup> fidelity or to determine how both routes evolved, but bioinformatic and phylogenetic analyses have led to a better understanding of the Thg1 evolution.

The Thg1 resemblance to polymerases indicates that Thg1 evolved from duplication of a polymerase palm domain and was then selected to perform reverse direction (3'→5') ligation of the G<sup>-1</sup> to tRNA<sup>His</sup>ΔG<sup>-1</sup>. Phylogenetic analysis indicates that Thg1 evolved in the common ancestor of archaea and eukarya and that the enzyme was only later horizontally transferred to certain bacterial lineages (Figure 2) (29). Since both the archaeal and bacterial tRNA<sup>His</sup> typically contain the C<sup>73</sup> discriminator base, the change to an A<sup>73</sup> discriminator base, which is ubiquitous in eukarya, must have occurred during the early evolution of eukarya. As mentioned above, when the discriminator base is changed to C the yeast Thg1 engages in prolonged reverse polymerization (37), and this undesired activity is inhibited when the A<sup>73</sup> is present. These data show a potential selective advantage for A<sup>73</sup> maintenance in eukaryal tRNA<sup>His</sup>.

Thg1, therefore, most likely evolved in the context of a tRNA<sup>His</sup> with a C<sup>73</sup> discriminator. The prerequisite for the altered RNase P cleavage is a C-discriminator base, which allows base pairing with the G<sup>-1</sup>-residue and an acceptor stem extension. At a time before Thg1 evolved, therefore, the G<sup>-1</sup>:C<sup>73</sup> base pair would have allowed facile aberrant RNase P cleavage (leaving G<sup>-1</sup> intact) to be a viable route for active tRNA<sup>His</sup> production. This ancestral situation is still retained in most bacterial lineages. An indication for the derivation of Thg1 from polymerases is also found in the biochemical data available from the archaeal-type Thg1. In cases where the selective pressure for a highly specific Thg1 reaction is low, Thg1-like proteins tend to display a more promiscuous enzymatic activity. In most bacteria and archaea, the G<sup>-1</sup> is genome encoded and thus Thg1 is not absolutely required for tRNA<sup>His</sup> fidelity. The same applies to *D. discoideum*, which contains four Thg1 homologs. In these cases, Thg1 activity is not as specific in its enzymatic activity as shown for the yeast enzyme, but is extended to tRNA repair functions, addition of other nucleotides and prolonged reverse polymerization. Nevertheless, these homologs are restricted in their ability to extend the reverse polymerization and preferentially add G or C residues (29,30,41).

## SUMMARY

During the past decade, the thorough Thg1 characterizations allow a clear understanding of how the guanylation reaction occurs. Taken together, the biochemical and structural data suggest that the reaction likely occurs in the following order of events (Figure 3): (I) as an initial step tRNA<sup>His</sup> anticodon recognition is carried out, possibly required for correct tRNA positioning and for directing the acceptor stem into the Thg1 active site. (II) in the active site, the tRNA then undergoes an activation step with ATP or GTP. After tRNA activation, the discriminator base allows correct tRNA positioning to lead to (III) the guanylation reaction with associated pyrophosphatase activity (IV).

This mechanistic overview is supported by data obtained from the variety of Thg1 assays reviewed herein. Step (I) occurs likely before Step (II), since

anticodon recognition is only of importance on unactivated tRNA substrates and can be circumvented by supplying ppp-tRNA substrates (V) (29,37,44). Thus, after the tRNA is guided into the active site, interaction with the anticodon must be less strong. Step (II) proceeds even in the presence of the “wrong” discriminator base and the activated tRNA intermediate can be trapped (34). At this point in the presence of the wrong discriminator base (as shown for archaeal enzymes acting on yeast tRNA<sup>His</sup> with an A<sup>73</sup>), the reaction stalls and the guanylation reaction is not completed (34). Thus, discriminator base positioning and/or recognition must occur between activation (II) and the guanylation Step (III). At what stage ATP and GTP are bound by the enzyme is unclear. ATP binding could occur prior to Steps (I) or (II), whereas GTP binding can possibly occur until just prior to (III). The obtained crystal structure with a dGTP in the potential ATP-binding pocket suggests, however, that tRNA binding is not a necessary prerequisite to nucleotide binding in this nucleotide-binding pocket. Also, the second nucleotide-binding pocket in the human Thg1 structure appears to be occupied by a nucleotide, even though specific interactions seem to be lacking. Possibly, GTP can enter the active site without tRNA being bound, but tight coordination and positioning might require the conformational changes induced by tRNA binding. As a final step, Thg1 associated pyrophosphatase activity (IV) leads to the PPI release and the mature tRNA<sup>His</sup> formation. This step must occur after the guanylation reaction is completed, since a small ppp-tRNA<sup>His</sup> amount has been observed in *in vitro* assays (20).

## OUTLOOK

The exciting advances in the Thg1 characterization made over the past decade signal that more discoveries lay ahead. Some challenging questions will have to be addressed to further characterize the *in vivo* Thg1 function in a number of organisms. Basic questions, such as where and how Thg1 binds to its substrate tRNA, could be addressed by a co-crystal structure of Thg1 with tRNA<sup>His</sup>. As a relatively small enzyme, Thg1 has to accomplish a number of tasks: anticodon and acceptor stem recognition, and cofactor and substrate binding. Whether each monomer would be able to simultaneously perform all these functions, or whether the functions are distributed between the subunits in a multimer remains to be elucidated.

In eukarya, the *in vivo* function and reaction mechanism have been studied in detail. The Thg1 deletion can be compensated by overexpression of HisRS and tRNA<sup>His</sup> $\Delta$ G<sup>-1</sup>, making extended essential Thg1 functions unlikely (27). Nevertheless, one study showed that Thg1 interacts with the multi subunit DNA-binding complex that binds to the origin of replication, the so termed origin recognition complex. Thus, a function in nuclear division and cytokinesis in yeast has been suggested (47). The question concerning why some bacteria and archaea encode Thg1 despite no apparent reason for its necessity,

is yet open for investigation. Thg1 may serve as a backup mechanism for tRNA quality control in case RNase P fails to properly process pre-tRNA<sup>His</sup>, as demonstrated for mitochondrial tRNA<sup>His</sup> (25). Some Thg1 homologs may have evolved entirely new functions in tRNA editing as suggested for the archaeal-type *D. discoideum* Thg1 homologs (41). In any case, both *in vitro* experiments and *in vivo* observations will ultimately determine what other roles Thg1 homologs might play in the cell.

Eukaryal- and bacterial-type Thg1 display a number of mechanistic differences, making Thg1 a potential drug target. The low abundance in bacteria would allow for a very specific targeting of organisms that carry a Thg1 homolog. For this purpose it is essential to investigate what the exact function of Thg1 in bacteria is, and whether it is an essential for survival or not.

Thg1 will become a future target of protein engineering efforts. The ability of the yeast enzyme to catalyze polymerization of G and C residues on mutant tRNA substrates, and the characterization of more promiscuous archaeal-type Thg1 enzymes indicates that Thg1 has potential to be selectively engineered using directed evolution techniques. An efficient 3'-5' RNA polymerase capable of carrying out a templated reaction would be a useful tool in molecular biology. Further, we could envision many applications for engineered Thg1 variants that specifically label, e.g. with a fluorescent probe, the 5'-end of a targeted RNA substrate.

## ACKNOWLEDGMENTS

We thank Laure Prat and Jiqiang Ling for helpful discussions and Emma Heinemann and Till O'Donoghue for encouragement.

## FUNDING

National Institute of General Medical Sciences (GM22854) and the National Science Foundation (MCB-0950474) (to D.S.). I.U.H. is a Postdoctoral fellow of the Deutsche Forschungsgemeinschaft (HE5802/1-1). Funding for open access charge: National Institute of General Medical Sciences (GM22854).

*Conflict of interest statement.* None declared.

## REFERENCES

- Giegé, R., Sissler, M. and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.*, **26**, 5017–5035.
- 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<sup>His</sup> in aminoacylation. *Nucleic Acids Res.*, **17**, 7855–7863.
- Francklyn, C. and Schimmel, P. (1990) Enzymatic aminoacylation of an eight-base-pair microhelix with histidine. *Proc. Natl Acad. Sci. USA*, **87**, 8655–8659.
- Rudinger, J., Florentz, C. and Giegé, R. (1994) Histidylolation 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.

5. Nameki, N., Asahara, H., Shimizu, M., Okada, N. and Himeno, H. (1995) Identity elements of *Saccharomyces cerevisiae* tRNA<sup>His</sup>. *Nucleic Acids Res.*, **23**, 389–394.
6. Fromant, M., Plateau, P. and Blanquet, S. (2000) Function of the extra 5'-phosphate carried by histidine tRNA. *Biochemistry*, **39**, 4062–4067.
7. Connolly, S.A., Rosen, A.E., Musier-Forsyth, K. and Francklyn, C.S. (2004) G-1:C73 recognition by an arginine cluster in the active site of *Escherichia coli* histidyl-tRNA synthetase. *Biochemistry*, **43**, 962–969.
8. Rosen, A.E., Brooks, B.S., Guth, E., Francklyn, C.S. and Musier-Forsyth, K. (2006) Evolutionary conservation of a functionally important backbone phosphate group critical for aminoacylation of histidine tRNAs. *RNA*, **12**, 1315–1322.
9. Rosen, A.E. and Musier-Forsyth, K. (2004) Recognition of G-1:C73 atomic groups by *Escherichia coli* histidyl-tRNA synthetase. *J. Am. Chem. Soc.*, **126**, 64–65.
10. Orellana, O., Cooley, L. and Söll, D. (1986) The additional guanylate at the 5' terminus of *Escherichia coli* tRNA<sup>His</sup> is the result of unusual processing by RNase P. *Mol. Cell. Biol.*, **6**, 525–529.
11. Burkard, U., Willis, I. and Söll, D. (1988) Processing of histidine transfer RNA precursors. Abnormal cleavage site for RNase P. *J. Biol. Chem.*, **263**, 2447–2451.
12. Brannvall, M., Fredrik Pettersson, B.M. and Kirsebom, L.A. (2002) The residue immediately upstream of the RNase P cleavage site is a positive determinant. *Biochimie.*, **84**, 693–703.
13. Ardell, D.H. and Andersson, S.G. (2006) TFAM detects co-evolution of tRNA identity rules with lateral transfer of histidyl-tRNA synthetase. *Nucleic Acids Res.*, **34**, 893–904.
14. Wang, C., Sobral, B.W. and Williams, K.P. (2007) Loss of a universal tRNA feature. *J. Bacteriol.*, **189**, 1954–1962.
15. Yuan, J., Gogakos, T., Babina, A.M., Söll, D. and Randau, L. (2011) Change of tRNA identity leads to a divergent orthogonal histidyl-tRNA synthetase/ tRNA<sup>His</sup> pair. *Nucleic Acids Res.*, **39**, 2286–2293.
16. Cooley, L., Appel, B. and Söll, D. (1982) Post-transcriptional nucleotide addition is responsible for the formation of the 5' terminus of histidine tRNA. *Proc. Natl Acad. Sci. USA*, **79**, 6475–6479.
17. Williams, J.B., Cooley, L. and Söll, D. (1990) Enzymatic addition of guanylate to histidine transfer RNA. *Methods Enzymol.*, **181**, 451–462.
18. Pande, S., Jahn, D. and Söll, D. (1991) Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae*. I. Purification and physical properties. *J. Biol. Chem.*, **266**, 22826–22831.
19. Jahn, D. and Pande, S. (1991) Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae*. II. Catalytic mechanism. *J. Biol. Chem.*, **266**, 22832–22836.
20. Gu, W., Jackman, J.E., Lohan, A.J., Gray, M.W. and Phizicky, E.M. (2003) tRNA<sup>His</sup> maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5' end of tRNA<sup>His</sup>. *Genes Dev.*, **17**, 2889–2901.
21. Heinemann, I.U., O'Donoghue, P., Madinger, C., Benner, J., Randau, L., Noren, C.J. and Söll, D. (2009) The appearance of pyrrolysine in tRNA<sup>His</sup> guanylyltransferase by neutral evolution. *Proc. Natl Acad. Sci. USA*, **106**, 21103–21108.
22. Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K. and Ohyama, K. (1992) Transfer RNA genes in the mitochondrial genome from a liverwort, *Marchantia polymorpha*: the absence of chloroplast-like tRNAs. *Nucleic Acids Res.*, **20**, 3773–3777.
23. Maréchal-Drouard, L., Kumar, R., Remacle, C. and Small, I. (1996) RNA editing of larch mitochondrial tRNA<sup>His</sup> precursors is a prerequisite for processing. *Nucleic Acids Res.*, **24**, 3229–3234.
24. Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.*, **15**, 57–61.
25. Placido, A., Sieber, F., Gobert, A., Gallerani, R., Giegé, P. and Maréchal-Drouard, L. (2010) Plant mitochondria use two pathways for the biogenesis of tRNA<sup>His</sup>. *Nucleic Acids Res.*, **38**, 7711–7717.
26. Gu, W., Hurto, R.L., Hopper, A.K., Grayhack, E.J. and Phizicky, E.M. (2005) Depletion of *Saccharomyces cerevisiae* tRNA<sup>His</sup> guanylyltransferase Thg1p leads to uncharged tRNA<sup>His</sup> with additional m<sup>5</sup>C. *Mol. Cell. Biol.*, **25**, 8191–8201.
27. Preston, M.A. and Phizicky, E.M. (2010) The requirement for the highly conserved G<sup>-1</sup> residue of *Saccharomyces cerevisiae* tRNA<sup>His</sup> can be circumvented by overexpression of tRNA<sup>His</sup> and its synthetase. *RNA*, **16**, 1068–1077.
28. Jackman, J.E. and Phizicky, E.M. (2006) tRNA<sup>His</sup> guanylyltransferase adds G<sup>-1</sup> to the 5' end of tRNA<sup>His</sup> by recognition of the anticodon, one of several features unexpectedly shared with tRNA synthetases. *RNA*, **12**, 1007–1014.
29. Heinemann, I.U., Randau, L., Tomko, R.J. Jr and Söll, D. (2010) 3'-5' tRNA<sup>His</sup> guanylyltransferase in bacteria. *FEBS Lett.*, **584**, 3567–3572.
30. Rao, B.S., Maris, E.L. and Jackman, J.E. (2011) tRNA 5'-end repair activities of tRNA<sup>His</sup> guanylyltransferase (Thg1)-like proteins from bacteria and archaea. *Nucleic Acids Res.*, **39**, 1833–1842.
31. Li, Y. and Zhou, H. (2009) tRNAs as regulators in gene expression. *Sci. China C Life Sci.*, **52**, 245–252.
32. Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J. and Hinnebusch, A.G. (2000) Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. *Mol. Cell*, **6**, 269–279.
33. Mascarenhas, C., Edwards-Ingram, L.C., Zeef, L., Shenton, D., Ashe, M.P. and Grant, C.M. (2008) Gen4 is required for the response to peroxide stress in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell.*, **19**, 2995–3007.
34. Abad, M.G., Rao, B.S. and Jackman, J.E. (2010) Template-dependent 3'-5' nucleotide addition is a shared feature of tRNA<sup>His</sup> guanylyltransferase enzymes from multiple domains of life. *Proc. Natl Acad. Sci. USA*, **107**, 674–679.
35. Francklyn, C., Shi, J.P. and Schimmel, P. (1992) Overlapping nucleotide determinants for specific aminoacylation of RNA microhelices. *Science*, **255**, 1121–1125.
36. Crothers, D.M., Seno, T. and Söll, D. (1972) Is there a discriminator site in transfer RNA? *Proc. Natl Acad. Sci. USA*, **69**, 3063–3067.
37. Jackman, J.E. and Phizicky, E.M. (2006) tRNA<sup>His</sup> guanylyltransferase catalyzes a 3'-5' polymerization reaction that is distinct from G<sup>-1</sup> addition. *Proc. Natl Acad. Sci. USA*, **103**, 8640–8645.
38. Hyde, S.J., Eckenroth, B.E., Smith, B.A., Eberley, W.A., Heintz, N.H., Jackman, J.E. and Doublé, S. (2010) tRNA<sup>His</sup> guanylyltransferase (THG1), a unique 3'-5' nucleotidyl transferase, shares unexpected structural homology with canonical 5'-3' DNA polymerases. *Proc. Natl Acad. Sci. USA*, **107**, 20305–20310.
39. Celesnik, H., Deana, A. and Belasco, J.G. (2007) Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol. Cell*, **27**, 79–90.
40. Deana, A., Celesnik, H. and Belasco, J.G. (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature*, **451**, 355–358.
41. Abad, M.G., Long, Y., Willcox, A., Gott, J.M., Gray, M.W. and Jackman, J.E. (2011) A role for tRNA<sup>His</sup> guanylyltransferase (Thg1)-like proteins from *Dictyostelium discoideum* in mitochondrial 5'-tRNA editing. *RNA*, **17**, 613–623.
42. 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.
43. Bullerwell, C.E. and Gray, M.W. (2005) In vitro characterization of a tRNA editing activity in the mitochondria of *Spizellomyces punctatus*, a Chytridiomycete fungus. *J. Biol. Chem.*, **280**, 2463–2470.
44. Perona, J.J. and Oza, J.P. (2010) Crystal structure of a reverse polymerase. *Proc. Natl Acad. Sci. USA*, **107**, 20149–20150.

45. Anantharaman,V., Iyer,L.M. and Aravind,L. (2010) Presence of a classical RRM-fold palm domain in Thg1-type 3'-5' nucleic acid polymerases and the origin of the GGDEF and CRISPR polymerase domains. *Biol. Direct.*, **5**, 43.
46. Jackman,J.E. and Phizicky,E.M. (2008) Identification of critical residues for G<sup>-1</sup> addition and substrate recognition by tRNA<sup>His</sup> guanylyltransferase. *Biochemistry*, **47**, 4817–4825.
47. Rice,T.S., Ding,M., Pederson,D.S. and Heintz,N.H. (2005) The highly conserved tRNA<sup>His</sup> guanylyltransferase Thg1p interacts with the origin recognition complex and is required for the G2/M phase transition in the yeast *Saccharomyces cerevisiae*. *Eukaryot. Cell.*, **4**, 832–835.