tRNA^{His} guanylyltransferase (THG1), a unique 3'-5' nucleotidyl transferase, shares unexpected structural homology with canonical 5'-3' DNA polymerases

Samantha J. Hyde^a, Brian E. Eckenroth^a, Brian A. Smith^{b,c}, William A. Eberley^b, Nicholas H. Heintz^{a,d}, Jane E. Jackman^{b,c,2}, and Sylvie Doublié^{a,1}

^aDepartment of Microbiology and Molecular Genetics, University of Vermont, Burlington, Vermont 05405; ^bDepartment of Biochemistry and Center for RNA Biology, and ^cOhio State Biochemistry Program, Ohio State University, 484 West 12th Avenue, Columbus, OH 43210; and ^dDepartment of Pathology, University of Vermont, Burlington, Vermont 05405

Edited by Olke C. Uhlenbeck, Northwestern University, Evanston, IL, and approved October 6, 2010 (received for review July 18, 2010)

All known DNA and RNA polymerases catalyze the formation of phosphodiester bonds in a 5' to 3' direction, suggesting this property is a fundamental feature of maintaining and dispersing genetic information. The tRNA^{His} guanylyltransferase (Thg1) is a member of a unique enzyme family whose members catalyze an unprecedented reaction in biology: 3'-5' addition of nucleotides to nucleic acid substrates. The 2.3-Å crystal structure of human THG1 (hTHG1) reported here shows that, despite the lack of sequence similarity, hTHG1 shares unexpected structural homology with canonical 5'-3' DNA polymerases and adenylyl/guanylyl cyclases, two enzyme families known to use a two-metal-ion mechanism for catalysis. The ability of the same structural architecture to catalyze both 5'-3' and 3'-5' reactions raises important questions concerning selection of the 5'-3' mechanism during the evolution of nucleotide polymerases.

G-1 addition | reverse polymerase | tRNA modification

All nucleotide polymerases, including DNA and RNA polymerases, reverse transcriptase, and telomerase, catalyze nucleotide addition in the 5' to 3' direction. The reaction involves the nucleophilic attack of a polynucleotide terminal 3'-OH onto the α -phosphate of an incoming nucleotide, followed by release of the pyrophosphate moiety. Although the 5' to 3' direction has been adopted by all polymerases and transferases described to date, there is one notable exception: the enzyme tRNA^{His} guanylyltransferase (Thg1). Thg1 catalyzes the highly unusual 3'-5' addition of a single guanine to the 5'-end of tRNA^{His} (1, 2). This reaction is an obligatory step in the maturation of this tRNA because the extra 5' base, G₋₁, constitutes a primary identity element for the aminoacyl-tRNA synthetase (HisRS) that attaches the amino acid histidine to the 3'-end of the tRNA (3–9). Thg1 is thus essential for maintaining the fidelity of protein synthesis. Consistent with the critical nature of the G₋₁ residue, THG1 is an essential gene in yeast and RNAi-mediated silencing of the Thg1 homolog in human cells results in severe cell-cycle progression and growth defects (2, 10, 11). Thg1 is widely conserved throughout eukarya, and Thg1 homologs are present in many archaea and bacteria.

In eukarya, G₋₁ addition occurs opposite a universally conserved A₇₃ and thus is the result of a nontemplated 3'-5' addition reaction. In addition, yeast Thg1 catalyzes a second reaction in vitro, extending tRNA substrates in the 3'-5' direction in a template-directed manner driven by Watson–Crick pairing (12). Thg1 enzymes in archaea also catalyze template-dependent 3'-5' addition, but do not catalyze nontemplated G₋₁ addition (13), suggesting that the templated 3'-5' addition reaction likely represents an ancestral activity of the earliest Thg1 family members.

The 3'-5' addition of G₋₁ to tRNA^{His} occurs via three chemical reactions, all catalyzed by Thg1 (2, 14) (Fig. 1). First, the 5'-monophosphorylated tRNA that results from RNase P cleavage of pre-tRNA^{His} is activated using ATP, creating a 5'-adenylylated-

Fig. 1. Thg1 catalytic steps. Thg1 catalyzes 3'-5' addition of G₋₁ to tRNA^{His} in three steps: adenylation, nucleotidyl transfer, and pyrophosphate removal.

tRNA^{His} intermediate. This adenylation step mirrors the activation step in aminoacyl-tRNA synthetases in which the amino acid receives an AMP moiety prior to being charged on the cognate tRNA (15, 16). In the second step, the 3'-hydroxyl of GTP attacks the activated intermediate, yielding triphosphorylated

Author contributions: J.E.J. and S.D. designed research; S.J.H., B.E.E., B.A.S., W.A.E., J.E.J., and S.D. performed research; N.H.H. contributed new reagents/analytic tools; S.J.H., B.E.E., B.A.S., J.E.J., and S.D. analyzed data; and J.E.J. and S.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The coordinates and structure factor amplitudes have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 3OTB (hTHG1-dGTP complex), 3OTC (Native II; unliganded hTHG1, trigonal form), 3OTD (iodide derivative), and 3OTE (Native I; unliganded hTHG1)].

See Commentary on page 20149.

¹To whom correspondence may be addressed at: E314 Given Building, 89 Beaumont Avenue, University of Vermont, Burlington, VT 05405. E-mail: sdoublié@uvm.edu.

²To whom correspondence may be addressed at: Department of Biochemistry, 484 West 12th Avenue, Columbus, OH 43210. E-mail: jackman.14@osu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1010436107/-DCSupplemental.

(ppp) G_{-1} -tRNA^{His}. Finally, the 5' pyrophosphate is removed, yielding mature, monophosphorylated (p) G_{-1} -containing tRNA^{His}. Although these chemical steps are reminiscent of activities catalyzed by well-studied RNA or DNA ligases or mRNA capping guanylyltransferase (17), no obvious sequence similarity exists between Thg1 and these or any other known enzyme families to suggest a possible molecular mechanism for Thg1 catalysis.

Thg1 is the only known example of an enzyme that catalyzes templated nucleotide addition in the 3'-5' direction, opposite to that of all known DNA and RNA polymerases. Thus the molecular mechanism of this enzyme is of great interest and structural characterization of Thg1 is essential for understanding this likely unique enzymology. We report here the structure of a eukaryal Thg1 family member determined at 2.3-Å resolution. The structure of human THG1 (hTHG1) reveals a shared active site architecture with canonical 5'-3' DNA polymerases and adenyl/ guanylyl cyclases, two families of enzymes that use a similar two-metal-ion mechanism for catalysis. Analysis of a crystal structure of hTHG1 bound to nucleotide and magnesium suggests that a preadenylation complex may have been captured.

Results

Overall Structure of Human THG1. The 2.3-Å crystal structure of hTHG1 represents a previously undetermined structure of an enzyme catalyzing 3'-5' nucleotide additions (Table S1). The hTHG1 construct used for structural studies was composed of 269 amino acids with a calculated molecular weight of 32 kDa. The purified protein eluted from gel exclusion chromatography with an apparent molecular weight of ~165 kDa (Fig. S1), consistent with formation of a higher order multimer in solution and with the tetrameric form of the enzyme observed in the crystals (Fig. 2). The tetrameric form of hTHG1 appears to be highly conserved, because yeast, archaeal, and bacterial Thg1 enzymes similarly eluted from gel exclusion with a molecular weight con-

sistent with a tetramer. hTHG1 was crystallized in two different crystal systems, tetragonal and trigonal, and in both cases the homotetramer appears as a dimer of dimers (Fig. 2), in which the first dimer constitutes the crystal asymmetric unit and the second is generated by symmetry.

The hTHG1 monomer is made of a β -sheet composed of six antiparallel strands flanked by three or four α -helices on each side. In addition, two antiparallel β -strands (β_6 and β_7) form a long arm that is seen only in the trigonal crystals. Interactions between the two monomers in the asymmetric unit are largely mediated by residues from helix α_D and β_4 . Several hydrogen bond interactions are made between main-chain atoms of β_4 and side-chain atoms of α_D , an example of which is seen between G129-N and T98-OH. Alteration of T98 to alanine disrupted multimer formation, as judged by gel exclusion chromatography (Fig. S1), and yielded a variant deficient in G_{-1} activity (Table 1). In addition to hydrogen bonding interactions, two salt bridges (K95_A-D128_B; E13_A-R130_B) stabilize the interface between the two monomers. The importance of these residues is underscored by mutational studies of *Saccharomyces cerevisiae* Thg1 (ScThg1), with which hTHG1 shares 52% sequence identity: Alteration of any of these four strictly conserved residues in yeast (ScK96, ScD131, ScE13, or ScR133) to alanine strongly diminishes G_{-1} addition activity (18).

The N-terminal segment comprising the first 20 residues from one monomer in the A/B dimer wraps around the other monomer in such a way that residues in helix α_A contact the nucleotide binding site of the other monomer (see below). The interface between the two monomers buries 4,200 Å² (19). The intertwined N-terminal segments from the first dimer provide a platform to interact with their symmetry-related counterparts from the second dimer. The dimer/dimer (AB/A'B') interface is less extensive than the monomer/monomer (A/B) interface, burying only 1,800 Å² (19). The tetrameric form was calculated to be the most stable oligomeric form ($\Delta G = -68$ kcal/mol vs. -22 kcal/mol for the dimer) (20).

Unexpected Homology to DNA Polymerases. Because hTHG1 shares no significant sequence similarity with any known protein, it was unclear whether hTHG1 would have any structural homologs. Surprisingly, significant homology was found with guanylyl and adenyl cyclases [$\beta\alpha\beta\alpha\beta$ motif of C_{1a} domain; Z score = 8.0; Protein Data Bank (PDB) ID codes 2W01 (21) and 3E8A (22)] (23), along with the palm domain of several traditional polymerases including T7 DNA polymerase, a family A polymerase [Z score = 5.3, PDB ID code 1T7P (24)], and DNA polymerase II, a member of the B family [Z score = 5.9, PDB ID code 1Q8I, (25)]. The fold of the hTHG1 $\beta\alpha\beta\alpha\beta$ motif (residues 22-135) most closely matches that of the cyclases. However, the superposition of hTHG1 with the polymerases suggests that the Thg1

Fig. 2. Ribbon diagram of the human THG1 homotetramer. The tetramer consists of a dimer of dimers. Monomers are colored as follows: gray, monomer A; yellow, monomer B; blue, monomer A'; and green, monomer B'. The long arm composed of β -strands β_6 and β_7 is seen in its entirety only in the trigonal crystals. Disordered residues are indicated with spheres.

Table 1. G_{-1} addition activity of yeast and human Thg1 variants

hTHG1 variant	ScThg1 position	% specific activity hTHG1*	% specific activity ScThg1
Wild type	—	100	100
D29A	same	0.5	<0.1
D76A	D77	<0.2 [†]	<0.03 [‡]
E77A	E78	5.3	<0.02 [‡]
T98A	T99	<0.7 [†]	NT [§]
H152A	H155	<0.7 [†]	0.04 [‡]
K187A	K190	1.4	<0.04 [‡]
N198A	N201	<0.7 [†]	<0.05 [‡]

*Specific activity of each Thg1 variant for G_{-1} addition to 5'-³²P-labeled yeast tRNA^{His} assayed as described (18) and shown as percent relative to wild type.

[†]Upper limit based on no detectable G_{-1} product observed at highest concentration of variant tested.

[‡]Data from ref. 18.

[§]Not tested.

Fig. 3. The catalytic core of hTHG1 most closely resembles DNA polymerases of the A family. (A) Comparison of topology diagrams for hTHG1, adenylate cyclase (PDB ID code 3E8A) (22), and palm domains of T7 DNA polymerase (family A, PDB ID code 1T8E) (42), and RB69 gp43 (family B, PDB ID code 2OYQ) (43). The numbering of the β -strands is for the palm domain only and locations of the catalytic carboxylate residues are indicated by black diamonds. (B) Superposition of hTHG1 (residues 22–135, blue) with the palm domain of T7 DNA polymerase (residues 466–486, 608–698, gray) overlays catalytic carboxylates. The incoming nucleotide and metal ions from the T7 DNA polymerase complex are shown (1T7P) (24).

mechanism is closest to that of the A family polymerases (Fig. 3), based on the position of three highly conserved carboxylate residues (see below). [All DNA polymerases harbor these three carboxylates in the polymerase active site (26, 27) and members of the A and B families differ in the relative position of the three residues (Fig. 3A)]. After the hTHG1 structure was solved, Aravind and coworkers published a computational analysis proposing a model which is consistent with parts of our biochemical and structural data, including the potential involvement of three carboxylates and divalent metal ions in catalysis (see below) (28).

The core β -sheet of the palm domain of T7 DNA polymerase superimposes with the corresponding motif in hTHG1 with an rmsd of 1.8 Å. The superposition pinpointed three residues that correspond to the three catalytic carboxylates of all known DNA polymerases (Fig. 3B) (29). In contrast, adenylate and guanylate cyclases contain only two catalytic aspartates and a cysteine, alanine or glycine in lieu of the third carboxylate (21, 30). The highly conserved hTHG1 carboxylates D29, D76, and E77 correspond to D475, D654, and E655 in T7 DNA polymerase. Alteration of any of these three residues to alanine decreased G_{-1} addition activity by hTHG1 (Table 1) in a manner consistent with that expected for catalytically important residues. Whereas D29A and D76A hTHG1 variants exhibited >100-fold decreased specific activity compared with wild-type hTHG1, E77 exhibited a more modest (20-fold) decrease in activity similar to the relatively minor effect of the E655A alteration in T7 DNA polymerase (24). Interestingly, the alanine alteration at the analogous position in ScThg1 (ScE78) severely impacts ScThg1 activity (Table 1), suggesting species-specific differences in the roles of the highly conserved glutamate. The superposition of the three strictly conserved Thg1 carboxylates with those of T7 and their crucial role in catalysis strongly suggest that Thg1 unexpectedly uses the two-metal-ion mechanism of canonical 5'-3' polymerases (31–33).

Two Metal Ions in the hTHG1 Active Site. During the three-step nucleotide addition reaction, Thg1 first binds ATP for adenylation

then GTP for the nucleotidyl transfer reaction (Fig. 1). We obtained 2.95-Å diffraction data from a complex of hTHG1 with 2'-deoxy-GTP (dGTP) and Mg^{2+} (the enzyme can use dGTP as well as GTP, ref. 12). The resulting electron density maps showed clear density for one dGTP and one triphosphate moiety (see below) in each monomer of the asymmetric unit (Fig. 4). The dGTP is located in what is predicted to be the active site of the enzyme based on its structural homology to DNA polymerases: A superposition of the dGTP-bound hTHG1 structure with a T7 DNA polymerase complex with DNA and ddGTP overlays the two nucleotides (Fig. S2). In the hTHG1 cocrystal structure, the guanine base stacks against two conserved hydrophobic residues, A37 and F42. The Watson–Crick face of the base is within hydrogen bonding distance of two protein backbone atoms: O6 and N1 contact the D47 amide and A43 carbonyl, respectively. The 3'-OH of the deoxyribose makes a hydrogen bond with H34 while the face of the sugar is within van der Waals contact of the atoms forming the peptide bond between F33 and H34. The β - and γ -phosphates interact with a main-chain amide via a nonbridging oxygen (β -phosphate with H34 and γ -phosphate with N32), and all three phosphates coordinate at least one of the two metal ions. The rmsd between the dGTP-bound and unliganded structures is 0.45 Å based on an alignment length of 234 amino acids, suggesting no major structural difference between the two structures. There are, however, notable changes in side-chain orientation that occur around the active site.

The dGTP-bound crystal structure revealed two bound Mg^{2+} ions associated with the nucleoside triphosphate (Fig. 4B). The presence of these ions was confirmed using an anomalous difference Fourier map calculated with data acquired from manganese soaked crystals (Fig. S3). The distance between the two metal ions is 4.3 Å. As with family A DNA polymerases, the two hTHG1 aspartates (D29 and D76) coordinate the two divalent metal ions, whereas E77 points away from the metals. Metal B contacts non-

Fig. 4. Structure of the dGTP-bound form of hTHG1. (A) Overall view of the dimer in the asymmetric unit showing monomer A (gray) and monomer B (yellow) and the bound nucleotides and metals. (B) Close-up of the hTHG1 nucleotide binding site. The structure reveals a bound dGTP with two Mg^{2+} ions, A and B (purple spheres). Also present in the crystal is an additional triphosphate moiety and Mg^{2+} . A simulated annealing omit map contoured at 3 σ is shown in blue. The catalytic carboxylates Asp29, Asp76, and Glu77 are shown, as are additional residues involved in nucleotide binding and hTHG1 function.

bridging oxygens of all three phosphates of dGTP and the two catalytic aspartates D29 and D76. An interaction with the main-chain oxygen of G30 completes the octahedral coordination. Metal A interacts with the two aspartates and a nonbridging oxygen of the α -phosphate.

Nucleotide-Bound hTHG1 Structure. The three-step reaction catalyzed by Thg1 requires diverse interactions with multiple nucleotide substrates, including ATP during adenylation, GTP during nucleotidyl transfer, and the 5'-monophosphorylated and 5'-triphosphorylated ends of the tRNA for adenylation and pyrophosphate removal, respectively (Fig. 1). Thus, the bound dGTP visualized in the hTHG1 structure could reflect the position of any of these nucleotide species. To examine these possibilities, we altered H34 and S75, two hTHG1 residues that contact the bound dGTP (Fig. 4B), to alanine and used single-turnover kinetic assays to individually measure rates of adenylation (step 1) or nucleotidyl transfer (step 2) during G_{-1} addition. H34 is highly conserved (replaced only by S, T, or K in eukarya) and is positioned to accept a hydrogen bond from the 3'-hydroxyl of the bound dGTP. S75 is universally conserved in Thg1 enzymes from all domains of life, and the serine hydroxyl is ~ 4 Å away from the N7 of the guanine base.

Removal of a side chain from a residue that activates the 3'-hydroxyl nucleophile for catalysis is expected to substantially decrease the rate of nucleotidyl transfer. However, the maximal rate constants for adenylation (k_{aden}) and nucleotidyl transfer (k_{ntrans}) are only modestly decreased by alteration of H34 to alanine (Table 2 and Fig. S4), suggesting that H34 does not serve this role. Similarly, the S75A alteration did not significantly affect k_{aden} or k_{ntrans} , indicating that neither of these residues participates directly in the chemical steps for adenylation or nucleotidyl transfer. However, a 10-fold increase in the $K_{D,\text{app,ATP}}$ for the adenylation step of the reaction is observed for hTHG1 S75A, with a relatively smaller effect on the $K_{D,\text{app,GTP}}$ for nucleotidyl transfer. Although the placement of this particular ligand-bound structure along the reaction coordinate is uncertain at this early stage, these results suggest that the bound dGTP in the structure does not represent the position of the incoming GTP that is incorporated into the growing polynucleotide chain. Instead, the results suggest that the bound dGTP may reveal the position of the ATP that is used for the activation step (adenylation). This interpretation is further supported by the orientation of the bound dGTP nucleotide with its triphosphate moiety coordinating two metal ions and thus poised for chemistry to occur at the α -phosphate of the bound NTP, such as occurs during adenylation (Fig. 1). Moreover, the 3'-OH of the dGTP does not contact either metal ion and is not positioned for nucleophilic attack.

Notably, interactions observed between hTHG1 and the nucleotide base are not restricted to a guanine. The interaction of the Watson-Crick face of the base with main-chain atoms are such that the binding site could accommodate either a guanine or an adenine by alternating backbone carbonyl/amide contacts. The interaction between guanine N7 and S75 is also possible with either purine. Importantly, although ATP is pre-

ferred by hTHG1, GTP is able to substitute for ATP in the tRNA activation step.

A Second Active Site Triphosphate. As described above, the electron density maps revealed the presence of a second bound molecule in each monomer, a triphosphate. The presence of the triphosphate group implies that a second nucleotide was bound; the base and sugar moieties are not seen in the electron density map presumably because of the lack of specific interactions. This situation is reminiscent of that described for the class I 3'-CCA-adding enzymes (tRNA nucleotidyl transferases), where the base moiety of the nucleotide is not tightly bound in the absence of the tRNA substrate (34–36).

The triphosphate moiety interacts with several strictly conserved residues: the α -phosphate contacts R27 and R130, whereas the γ -phosphate contacts R130, as well as R92 and K95 of the adjacent monomer. A metal ion, presumably Mg^{2+} , is also bound by the α -, β -, and γ -phosphates. The density was assigned to a metal ion rather than a water molecule by analogy with other polymerase structures deposited in the Protein Data Bank in which the triphosphate tail invariably ligates a metal ion. Altering any of the aforementioned positively charged residues to alanine in ScThg1 causes a marked decrease in G_{-1} addition activity, suggesting that this site is critical for catalysis (18). Based on the data presented above, the second bound nucleotide is not likely to reflect the position of the ATP used for adenylation. Therefore, the observed triphosphate may pinpoint a binding site for the incoming nucleotide that participates in 3'-5' addition or the tRNA molecule (Fig. 1); further biochemical and structural characterization with bound tRNA and nucleotide substrates will be needed to evaluate these possibilities.

tRNA Recognition by hTHG1. The important question of how hTHG1 binds to substrate tRNA is not addressed by the current hTHG1 structure, because no known tRNA-binding structural motifs were identified. A role in tRNA-binding may help to explain the observation that several highly conserved residues (H152, K187, E189, N198, and K208) are critically important to catalysis (18) (Table 1), yet are located in a small helical subdomain (helices α F and α G) 20–30 Å from the nucleotide binding site (Fig. 4A). In addition, previous biochemical characterization of ScThg1 implicated a single aspartate residue (ScD68, analogous to hTHG1 D67) in tRNA^{His} anticodon recognition (18). Nonetheless, attempts to build a model of tRNA-bound hTHG1 suggest multiple possibilities for the orientation and mode of tRNA substrate binding, and additional biochemical and structural data are needed to critically evaluate these potential alternatives. Thus the exact binding mode of tRNA to Thg1 remains uncertain, and will have to await the crystal structure of the enzyme with its polynucleotide substrate.

Discussion

Despite sharing several biochemical features with aminoacyl-tRNA synthetases and DNA/RNA ligases (37), structural characterization of Thg1 has not revealed further similarities between Thg1 and either of these enzyme families. Instead, our structural data show that Thg1 shares a similar active site architecture with adenylyl/guanylyl cyclases and family A DNA polymerases and likely uses the well-characterized two-metal-ion mechanism (24, 30–32, 38, 39) for catalysis of 3'-5' nucleotide addition. This proposed mechanism is a previously undescribed example of use of the two-metal-ion active site for nucleotide addition in the 3'-5' direction (Fig. 5). Each step of Thg1 catalysis is essentially a phosphoryl transfer reaction involving nucleophilic attack on a high-energy phosphate bond (Fig. 1), the same chemistry catalyzed by other two-metal-ion dependent enzymes such as DNA and RNA polymerases that perform canonical 5'-3' nucleotide addition or by proofreading polymerases with 3'-5' exonuclease

Table 2. Kinetic analysis of residues interacting with the nucleotide

hTHG1	k_{aden} , min ⁻¹	$K_{D,\text{ATP}}$, μM	k_{ntrans} , min ⁻¹	$K_{D,\text{GTP}}$, μM
Wild type	0.58 ± 0.02	210 ± 20	0.097 ± 0.005	9.9 ± 4.9
S75A	0.72 ± 0.11	2600 ± 500	0.043 ± 0.008	27 ± 17
H34A	0.31 ± 0.01	130 ± 10	0.07^*	ND [†]

*Lower limit to k_{ntrans} based on measurement of k_{obs} at a single high [GTP] (250 μM).

[†]Not determined.

Fig. 5. Predicted two-metal-ion mechanism for 3'-5' nucleotide addition. Proposed mechanism for 5'-adenylation catalyzed by hTHG1 (A) based on structural analogy with 5'-3' nucleotide addition catalyzed by T7 DNA polymerase (B). For hTHG1, two additional ligands (X) are proposed to complete the expected octahedral coordination sphere for Me_A ; these are likely to be water molecules that cannot be positively identified at the resolution of the current structure. Although activation of the tRNA 5'-phosphate for attack on the ATP α -phosphate is not strictly required, this step may be enhanced by coordination of the tRNA 5'-phosphate to Me_A (red dashed line), by analogy to the similar role demonstrated for Me_A in nucleophile activation in T7 DNA polymerase.

activity. These results provide a glimpse into an unusual active site and raise many intriguing questions about the molecular basis for nucleotide selection during templated vs. nontemplated 3'-5' addition reactions, the nature of the rearrangements that necessarily occur to accommodate all three steps of the Thg1 reaction, and details of recognition of tRNA substrate, all of which will require further biochemical and structural characterization.

The ability of the same active site features to be used for both 5'-3' and 3'-5' addition suggested by our structural and biochem-

ical characterization raises questions about the origins of nucleotide addition enzymes, which are fundamental to biology (26, 40). If nucleotidyl transferases can use both 5'-3' and 3'-5' reactions, why did nature choose 5'-3' over 3'-5' addition activity for DNA and RNA polymerases? The 3'-5' addition reaction catalyzed by Thg1 requires consumption of an additional ATP to activate a monophosphorylated 5'-end for addition of the first nucleotide. However, Thg1 can use the triphosphorylated 5'-end generated after nucleotidyl transfer (Fig. 1) for subsequent nucleotide additions (12). Thus overall levels of ATP consumption differ little for polymerization using 5'-3' vs. 3'-5' addition reaction mechanisms and an argument based purely on cell energetics does not explain the predominance of 5'-3' addition in biology. Often the advantage in fidelity of replication offered by proofreading mechanisms is suggested to account for the predominance of the 5'-3' mechanism of canonical polymerases, because 3'-5' addition would require reactivation of the monophosphorylated tRNA 5'-end generated by excision of an incorrectly added nucleotide. The fact that Thg1 nonetheless catalyzes this reactivation step, apparently within the same active site, suggests that additional biochemical considerations or constraints may have led to the prevalence of 5'-3' nucleotidyl addition in biology.

Materials and Methods

Crystal Structure Determination. Tetragonal and trigonal crystals of hTHG1 were obtained by vapor diffusion and cryoprotected as described in the *SI Text*. X-ray data were collected at the Advance Photon Source and in-house (Table S1). The structure of hTHG1 was solved with a single iodide derivative using anomalous and isomorphous data. Other structures were solved by difference Fourier methods if isomorphous, or molecular replacement otherwise. COOT (41) was used for model building and crystallographic refinement was performed with Crystallography and NMR System 1.2 (19) (Table S1 and Fig. S5). Details are in *SI Text*.

Biochemical Assays. G_{-1} addition to 5'- ^{32}P -labeled yeast tRNA^{His} was assayed as described previously (18). Single-turnover kinetic assays were performed using either 5'- ^{32}P -labeled yeast tRNA^{His} in the presence of ATP (for adenylation) or γ - ^{32}P -labeled yeast ppp-tRNA^{His} in the presence of GTP (for nucleotidyl transfer). For detailed descriptions of methods employed for biochemical assays, see *SI Text*.

ACKNOWLEDGMENTS. We thank Dr. P. Aller and K. Zahn for collecting diffraction datasets at the Advance Photon Source synchrotron. We thank Dr. F. Faucher for help with improving crystals. We thank Dr. C. S. Francklyn for discussions and Drs. E. M. Phizicky, C. R. H. Raetz, and M. A. Rould for critically reading the manuscript. The work on Thg1 mechanism is supported by National Institutes of Health Grant GM087543 (to J.E.J.). The General Medicine and Cancer Institutes Collaborative Access Team has been funded in whole or in part with Federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under Contract DE-AC02-06CH11357.

- Cooley L, Appel B, 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.
- Gu W, Jackman JE, Lohan AJ, Gray MW, Phizicky EM (2003) tRNA^{His} maturation: An essential yeast protein catalyzes addition of a guanine nucleotide to the 5' end of tRNA^{His}. *Genes Dev* 17:2889–2901.
- Himeno H, et al. (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.
- Nameki N, Asahara H, Shimizu M, Okada N, Himeno H (1995) Identity elements of *Saccharomyces cerevisiae* tRNA(His). *Nucleic Acids Res* 23:389–394.
- Rosen AE, Brooks BS, Guth E, Francklyn CS, Musier-Forsyth K (2006) Evolutionary conservation of a functionally important backbone phosphate group critical for aminoacylation of histidine tRNAs. *RNA* 12:1315–1322.
- Rosen AE, Musier-Forsyth K (2004) Recognition of $G_{-1}:C73$ atomic groups by *Escherichia coli* histidyl-tRNA synthetase. *J Am Chem Soc* 126:64–65.
- Rudinger J, Florentz C, Giegé R (1994) Histidylated by yeast HisR 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.
- Yan W, Francklyn C (1994) Cytosine 73 is a discriminator nucleotide in vivo for histidyl-tRNA in *Escherichia coli*. *J Biol Chem* 269(13):10022–10027.
- Gu W, Hurto RL, Hopper AK, Grayhack EJ, Phizicky EM (2005) Depletion of *Saccharomyces cerevisiae* tRNA(His) guanylyltransferase Thg1p leads to uncharged tRNA^{His} with additional m⁵C. *Mol Cell Biol* 25:8191–8201.
- Rice TS, Ding M, Pederson DS, Heintz NH (2005) The highly conserved tRNA^{His} guanylyltransferase Thg1p interacts with the origin recognition complex and is required for the G2/M phase transition in the yeast *Saccharomyces cerevisiae*. *Eukaryotic Cell* 4:832–835.
- Guo D, et al. (2004) Identification and characterization of a novel cytoplasmic protein ICF45 that is involved in cell cycle regulation. *J Biol Chem* 279:53498–53505.
- Jackman JE, Phizicky EM (2006) tRNA^{His} guanylyltransferase catalyzes a 3'-5' polymerization reaction that is distinct from G_{-1} addition. *Proc Natl Acad Sci USA* 103:8640–8645.
- Abad MG, Rao BS, Jackman JE (2010) Template-dependent 3'-5' nucleotide addition is a shared feature of tRNA^{His} guanylyltransferase enzymes from multiple domains of life. *Proc Natl Acad Sci USA* 107:674–679.
- Jahn D, Pande S (1991) Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae*. II. Catalytic mechanism. *J Biol Chem* 266:22832–22836.

15. Ibba M, Söll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69:617–650.
16. Francklyn C, Perona JJ, Puetz J, Hou YM (2002) Aminoacyl-tRNA synthetases: Versatile players in the changing theater of translation. *RNA* 8:1363–1372.
17. Shuman S, Lima CD (2004) The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyl transferases. *Curr Opin Struct Biol* 14:757–764.
18. Jackman JE, Phizicky EM (2008) Identification of critical residues for G_{-1} addition and substrate recognition by tRNA(His) guanylyltransferase. *Biochemistry* 47:4817–4825.
19. Brunger AT (2007) Version 1.2 of the crystallography and NMR system. *Nat Protoc* 2:2728–2733.
20. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
21. Rauch A, Leipelt M, Russwurm M, Steegborn C (2008) Crystal structure of the guanylyl cyclase Cya2. *Proc Natl Acad Sci USA* 105:15720–15725.
22. Mou TC, Masada N, Cooper DM, Sprang SR (2009) Structural basis for inhibition of mammalian adenylyl cyclase by calcium. *Biochemistry* 48:3387–3397.
23. Holm L, Sander C (1997) Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res* 25:231–234.
24. Doublé S, Tabor S, Long AM, Richardson CC, Ellenberger T (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 391:251–258.
25. Wang F, Yang W (2009) Structural insight into translesion synthesis by DNA Pol II. *Cell* 139:1279–1289.
26. Ito J, Braithwaite DK (1991) Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* 19:4045–4057.
27. Delarue M, Poch O, Tordo N, Moras D, Argos P (1990) An attempt to unify the structure of polymerases. *Protein Eng* 3:461–467.
28. Anantharaman V, Iyer LM, 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.
29. Steitz TA (1999) DNA polymerases: Structural diversity and common mechanisms. *J Biol Chem* 274:17395–17398.
30. Tesmer JJ, et al. (1999) Two-metal-ion catalysis in adenylyl cyclase. *Science* 285:756–760.
31. Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci USA* 90:6498–6502.
32. Steitz TA (1998) A mechanism for all polymerases. *Nature* 391:231–232.
33. Doublé S, Sawaya MR, Ellenberger T (1999) An open and closed case for all polymerases. *Structure* 7:R31–35.
34. Xiong Y, Li F, Wang J, Weiner AM, Steitz TA (2003) Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol Cell* 12:1165–1172.
35. Xiong Y, Steitz TA (2006) A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr Opin Struct Biol* 16:12–17.
36. Betat H, Rammelt C, Morl M (2010) tRNA nucleotidyl transferases: Ancient catalysts with an unusual mechanism of polymerization. *Cell Mol Life Sci* 67:1447–1463.
37. Jackman JE, Phizicky EM (2006) tRNA(His) guanylyltransferase adds G_{-1} to the 5' end of tRNA(His) by recognition of the anticodon, one of several features unexpectedly shared with tRNA synthetases. *RNA* 12:1007–1014.
38. Beese LS, Steitz TA (1991) Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: A two metal ion mechanism. *EMBO J* 10:25–33.
39. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J (1994) Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* 264:1891–1903.
40. Filee J, Forterre P, Sen-Lin T, Laurent J (2002) Evolution of DNA polymerase families: Evidences for multiple gene exchange between cellular and viral proteins. *J Mol Evol* 54:763–773.
41. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
42. Briebe LG, et al. (2004) Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *EMBO J* 23:3452–3461.
43. Zahn KE, Belrhali H, Wallace SS, Doublé S (2007) Caught bending the A-rule: Crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry* 46:10551–10561.