Cell, Vol. 111, 815–824, December 13, 2002, Copyright ©2002 by Cell Press

Crystal Structures of the *Bacillus stearothermophilus* CCA-Adding Enzyme and Its Complexes with ATP or CTP

Fang Li,[1] Yong Xiong,[1] Jimin Wang,[1] HyunDae D. Cho,[4] Kozo Tomita,[4] Alan M. Weiner,[4] and Thomas A. Steitz[1,2,3,5]

[1]Department of Molecular Biophysics and Biochemistry
[2]Department of Chemistry
[3]Howard Hughes Medical Institute
Yale University
New Haven, Connecticut 06520
[4]Department of Biochemistry
School of Medicine
University of Washington
Seattle, Washington 98195

Summary

CCA-adding enzymes polymerize CCA onto the 3′ terminus of immature tRNAs without using a nucleic acid template. The 3.0 Å resolution crystal structures of the CCA-adding enzyme from *Bacillus stearothermophilus* and its complexes with ATP or CTP reveal a seahorse-shaped subunit consisting of four domains: head, neck, body, and tail. The head is structurally homologous to the palm domain of DNA polymerase β but has additional structural features and functions. The neck, body, and tail represent new protein folding motifs. The neck provides a specific template for the incoming ATP or CTP, whereas the body and tail may bind tRNA. Each subunit has one active site capable of switching its base specificity between ATP and CTP, an important component of the CCA-adding mechanism.

Introduction

Perhaps the most mysterious unsolved puzzle among the polynucleotide polymerases is how the CCA-adding enzyme can add a specific CCA sequence to the end of a tRNA without the use of a polynucleotide template. The CCA 3′ terminus found on all mature tRNAs is absolutely required for tRNA aminoacylation (Sprinzl and Cramer, 1979) and specifically interacts with the A and P loops of the ribosome peptidyltransferase center (Green and Noller, 1997; Nissen et al., 2000). The CCA-adding enzyme adds these three nucleotides with high specificity to tRNA nucleotide-73, using CTP and ATP as substrates. CCA-adding activity is highly conserved throughout evolution, and the activity has been identified in all three kingdoms (Yue et al., 1998). In all eukarya and archaea and in certain eubacteria, some or all tRNA genes do not encode the CCA sequence that is essential for the viability of the organisms (Aebi et al., 1990). In organisms whose tRNA genes encode the CCA sequence, mainly eubacteria, the CCA-adding enzyme maintains and repairs the tRNA 3′ CCA terminus, and inactivation of the gene encoding the CCA-adding enzyme impairs cell growth (Zhu and Deutscher, 1987). Here, we provide structural insights into some of the mechanisms by which this unique polymerase specificity is achieved.

Several different models have previously been proposed to explain how the CCA-adding enzymes can recognize appropriately either CTP or ATP in order to synthesize CCA. An early model suggested different binding sites for CTP and ATP on the single polypeptide (Deutscher, 1972, 1982; Tomari et al., 2000; Seth et al., 2002). However, there is only one domain that is homologous to the catalytic site of DNA polymerase β (Holm and Sander, 1995), and mutation of either of the two putative Mg^{2+} binding carboxylates abolishes the addition of both CTP and ATP, consistent with a common catalytic site for both nucleotides (Yue et al., 1998). A second model postulated a "collaborative templating" between the elongating 3′ terminus and a nucleotide binding site to alter the specificity at a single incorporation site (Shi et al., 1998). Finally, recognizing that the enzyme is an oligomer, it was suggested that one subunit might be specific for adding two C's with scrunching at the 3′ terminus, after which the 3′ terminus would shuttle to a second subunit that was specific for adding A (Li et al., 2000). Compounding the diversity of this family of enzymes and in some ways consistent with this "scrunching shuttling" model was the recent discovery that CCA addition in the eubacterium *Aquifex aeolicus* requires two different but related polypeptides, a CC-adding enzyme and an A-adding enzyme (Tomita and Weiner, 2001).

The CCA-adding enzyme is a member of the DNA polymerase β or nucleotidyltransferase (NT) superfamily (Holm and Sander, 1995), which has been divided into two classes on the basis of sequence similarity (Yue et al., 1998). Class I includes archaeal CCA-adding enzymes as well as DNA polymerase β (pol β), eukaryotic poly(A) polymerases (PAP), terminal deoxynucleotidyltransferases (TdT), and kanamycin nucleotidyltransferases (KNT). These enzymes share sequence similarity with one another only in the NT superfamily signature motif region. Class II includes both eubacterial and eukaryotic CCA-adding enzymes as well as eubacterial PAPs. Class II enzymes share a homologous 25 kDa N-terminal region, but differ from one another in their C-terminal regions.

To date, only three enzymes are known to be template-independent polymerases, and they all belong to the NT superfamily; these are the CCA-adding enzyme, PAP, and TdT. Of these three, the CCA-adding enzyme exhibits the greatest specificity by recognizing and incorporating both CTP and ATP in a defined CCA sequence. In contrast, TdT adds any of the nucleotides without selectivity while PAP only recognizes ATP and adds heterogeneous lengths of A's. Crystal structures have been solved for four members of class I enzymes: pol β, eukaryotic PAP, TdT, and KNT (Pelletier et al., 1994; Bard et al., 2000; Martin et al., 2000; Dalarue et al., 2002; Sakon et al., 1993). Unfortunately, neither of the two known structures of eukaryotic PAPs explained their nucleotide selection strategy. In the yeast PAP

[5]Correspondence: eatherton@csb.yale.edu

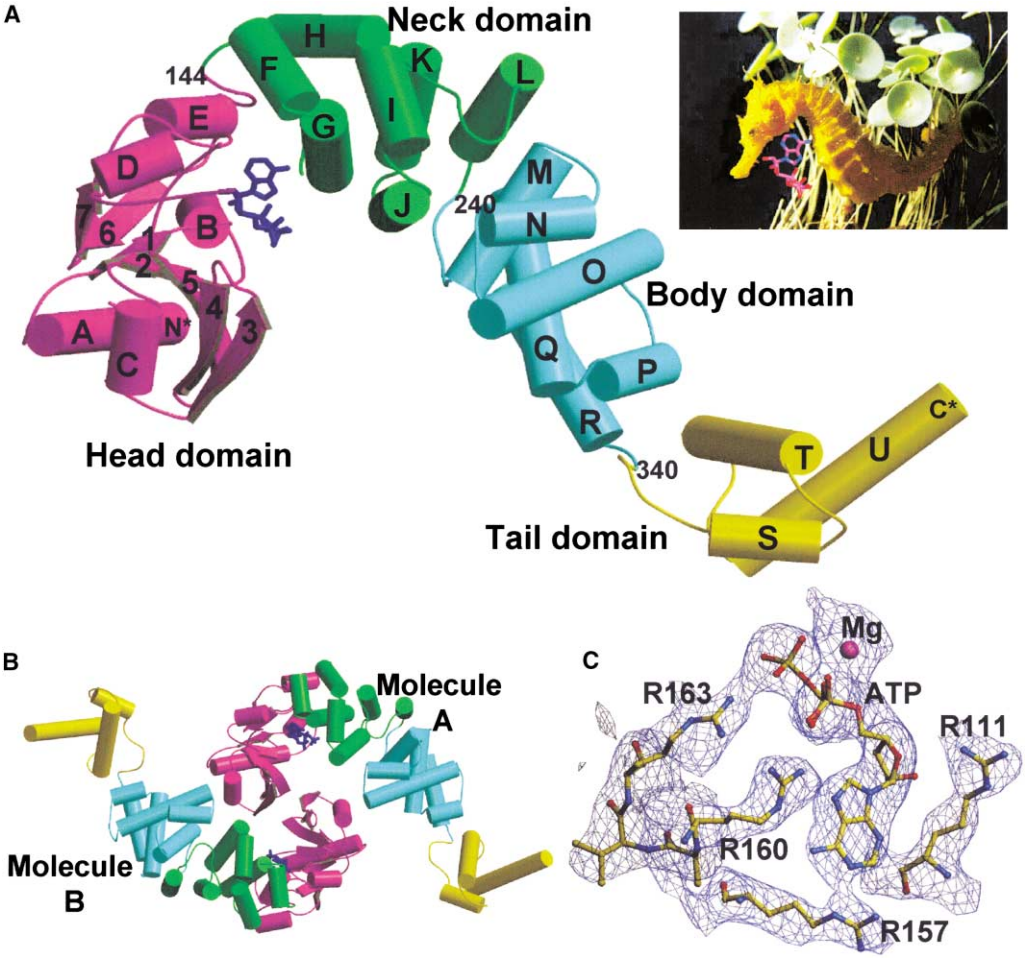

Figure 1. Overall Structure of BstCCA

(A) Structure of the seahorse-shaped BstCCA monomer. The head domain (residue 1–144) is shown in magenta, the neck domain (residue 145–240) in green, the body domain (residue 241–340) in cyan, and the tail domain (residue 341–404) in yellow. The bound ATP molecule is colored dark blue. Secondary structures are: strand 1 (β1), residues 22–25; β2, 42–45; β3, 59–62; β4, 69–72; β5, 76–81; β6, 118–121; β7, 124–127; helix A (αA), 3–17; αB, 27–34; αC, 49–55; αD, 104–108; αE, 132–138; αF, 146–153; αG, 155–168; αH, 174–183; αI, 184–190; αJ, 192–204; αK, 208–218; αL, 230–239; αM, 248–259; αN, 264–271; αO, 274–291; αP, 299–305; αQ, 307–321; αR, 325–338; αS, 350–357; αT, 363–378; and αU, 386–404.

(B) Structure of BstCCA dimer. Domains and ATP molecules are colored the same as in (A).

(C) Experimental electron-density map (1.0σ) of ATP and some surrounding protein residues. The densities were calculated from MAD phases, followed by solvent flattening and NCS averaging. All figures were generated using program Molscript (Kraulis, 1991).

structure (Bard et al., 2000), the base of ATP does not contact the enzyme even though the ATP is bound to the catalytic domain as occurs in pol β. The ATP observed in the bovine PAP structure is in an inactive conformation that is incompatible with the phosphoryl transfer reaction (Martin et al., 2000). How template-independent polymerases specifically select nucleotides remains to be elucidated. No structural information is available for class II enzymes or for archaeal CCA-adding enzymes of class I.

Here, we report the crystal structures of the unliganded CCA-adding enzyme from a eubacterium, *Bacillus stearothermophilus*, as well as its complexes with ATP or CTP, showing that a single active site can specifically recognize both ATP and CTP. These structures also provide insights into the mechanism of sequential CCA addition and have implications for the structures and functions of other class II enzymes.

Results and Discussion

Structure Determination

After screening both class I and class II CCA-adding enzymes from seven archeal and eubacterial species for crystal formation, we found that the full-length CCA-adding enzyme from *Bacillus stearothermophilus* (BstCCA) forms suitable crystals. Complexes with ATP or CTP were prepared by soaking the nucleotides into crystals. All three crystal forms are nearly isomorphous and contain two copies of BstCCA molecules per asymmetric unit related by a 2-fold noncrystallographic symmetry axis (Figures 1A and 1B). The structure of the

Table 1. Phasing and Refinement Statistics

Phasing Statistics for BstCCA/ATP				
Wavelength (Å)	1.1063	0.9786	0.9789	0.9560
Space group	$P3_121$ (a = b = 105.6 Å, c = 184.2 Å, $\alpha = \beta = 90°$, $\gamma = 120°$)			
Resolution (Å)	100–3.1	100–3.1	100–3.1	100–3.1
Phasing power iso / ano	1.22 / 1.04	0.94 / 1.39	– / 1.29	1.57 / 1.26
Figure of merit	0.50 (MAD); 0.59 (solvent flattening)			

Refinement Statistics	BstCCA/ATP	BstCCA/CTP	apo-BstCCA
Resolution (Å)	100–3.0	20–3.5	20–3.5
R_{sym} [a]	3.5%	4.8%	6.3%
Observed (unique) reflections	241,651 (24,280)	155,697 (15,352)	124,736 (15,392)
% complete (last shell)	99.8 (99.9)	99.8 (99.9)	99.4 (99.4)
I/σ (last shell)	41.6 (1.7)	32.6 (1.8)	19.0 (1.7)
R_{work} (R_{free}) [b]	22.8% (26.3%)	28.6% (33.1%)	28.8% (32.7%)

[a] $R_{sym} = \Sigma_{i,h}|I_{i,h} - \langle I_h \rangle|/\Sigma_{i,h}|I_{i,h}|$ where $\langle I_h \rangle$ is the mean of the i observations of the reflection h.
[b] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$. R_{free} is the same statistic, but calculated from a subset of the data (10%) that has not been used for refinement.

BstCCA complex with ATP was initially solved by multi-wavelength anomalous diffraction (MAD) using selenomethionine-labeled protein (Figure 1C). The structure was refined to an R factor (R_{free}) of 22.8% (26.3%). Subsequently, the model was refined against the data sets from crystals of a complex with CTP and of the apo-enzyme (Table 1).

Overall Architecture

The general architecture of BstCCA can be described as seahorse-shaped, with a head domain containing a seven-stranded β sheet and connecting helices, a right-handed superhelical neck domain, a loose helix-bundle body domain, and a left-handed superhelical tail domain (Figure 1A). A portion of the head domain is structurally homologous to the palm domain of pol β. The neck, body, and tail domains, however, represent new protein folding motifs; a search of the protein database for proteins with similar folds using the Dali server (Holm and Sander, 1998) did not reveal any previous structures that are convincingly related. The extended shape and domain arrangement of BstCCA molecule are different from all other polymerases with known structures. Both DNA polymerase I (pol I) structures and pol β structure have been likened to a hand with three domains: thumb, palm, and fingers (Ollis et al., 1985; Pelletier et al., 1994; Brautigam and Steitz, 1998). The corresponding domains have similar functions although the palms have different topologies. In pol β, the thumb domain interacts with the primer-template DNA substrate and the palm domain contains three carboxylate residues that bind the catalytically essential metal ions, while the fingers domain stabilizes the templating nucleic acid as well as the rest of the single-stranded DNA template (Figure 2A). In contrast, the thumb domain does not exist in BstCCA because the substrate is the single-stranded 3′ tRNA end. The head domain has a similar structure to the pol β palm domain for its N-terminal 80 residues and contains three carboxylates located in identical places (Figures 2A and 2B). Since BstCCA does not use a nucleic acid template, the neck domain of BstCCA, which occupies a similar position to the pol β fingers domain, provides the template for the incoming ATP or CTP (Figures 2A, 3, and 4). The extended body and tail domains of BstCCA may be involved in tRNA binding and they are absent in the pol β structure. The domain arrangement of BstCCA illustrates how the structure is dedicated to the special biological functions of CCA-adding enzymes.

The BstCCA enzyme appears to be a dimer in solution in the presence or absence of tRNA (F.L., unpublished) and exists as a dimer in the crystal whose subunits are related by a noncrystallographic 2-fold axis (Figure 1B). The surface area buried per subunit upon formation of this dimer is 512 Å^2, which is on the lower side of the range of surface areas observed to be buried in other dimeric proteins (Chothia and Janin, 1975). The distance between the two active sites measured between the Mn^{2+} bound to the active site carboxylates is 30 Å in a straight line and 42 Å following the molecular surface.

Structure and Function of the Head Domain

A portion of the head domain of BstCCA is structurally homologous to the palm domain of pol β, confirming that the eubacterial CCA-adding enzyme is indeed a member of the NT superfamily (Martin and Keller, 1996; Yue et al., 1998). The palm domain of BstCCA contains the five-stranded β sheet connected by two α helices that is also seen in pol β and other class I enzymes (Figures 2A and 2B; Pelletier et al., 1994; Bard et al., 2000; Martin et al., 2000; Delarue et al., 2002). The head domain of BstCCA, however, has two additional β strands to form a seven-stranded β sheet. In the region of BstCCA that is structurally homologous to pol β, 41 Cα atoms of BstCCA superimpose on the corresponding atoms of pol β with an rmsd of 1.7 Å. This alignment also superimposes three carboxylates of BstCCA (Asp40, Asp42, and Glu79) and the three carboxylates of pol β (Asp190, Asp192, and Asp256) that bind the catalytic Mg^{2+} ions (Figures 2A and 2B). Previous studies showed that Asp40 and Asp42 are part of the NT superfamily signature motif, and mutation of the corresponding residues in other NT enzymes abolishes polymerase activities (Martin and Keller, 1996). A third catalytic carboxylate in CCA-adding enzymes has long been sought unsuccessfully (Yue et al., 1998). Whether or not Glu79

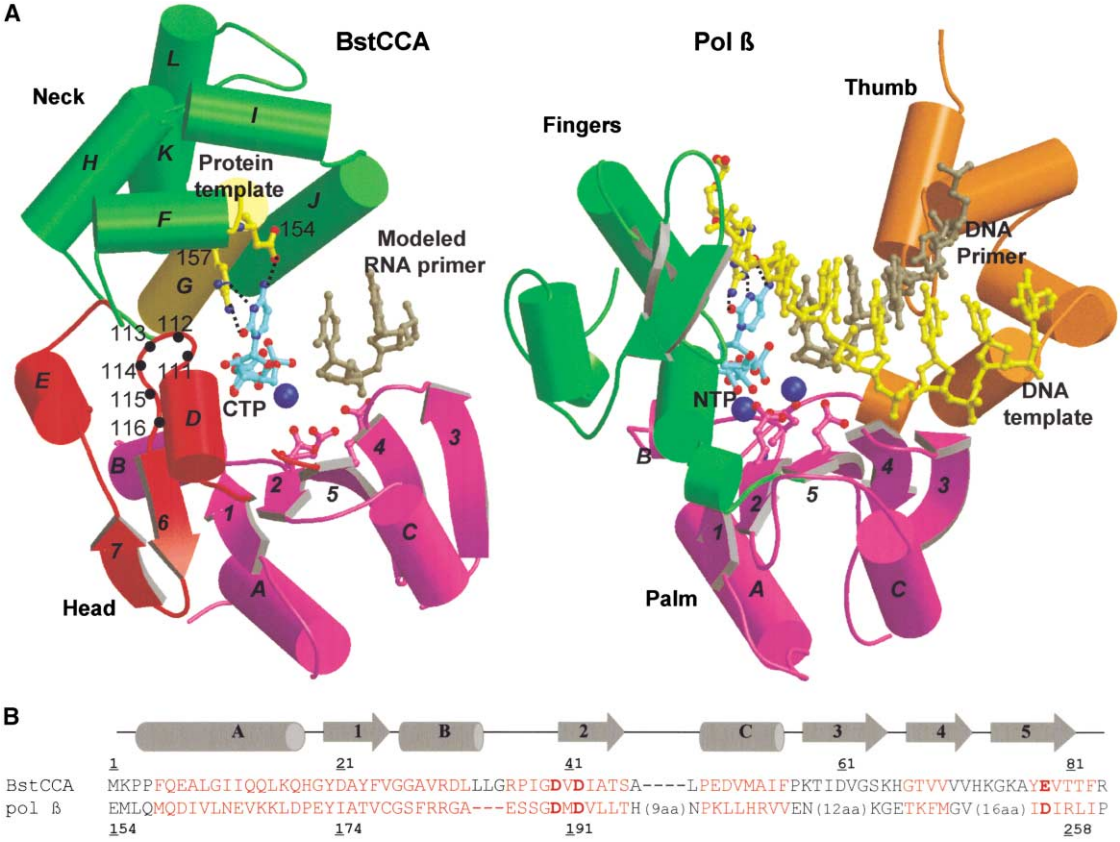

Figure 2. Structural Comparisons between BstCCA and pol β

(A) BstCCA polymerase site (head and neck domains) (left) and pol β polymerase site (thumb, palm, and fingers domains) (right). In the pol β structure (Pelletier et al., 1994), the thumb domain is shown in gold, the palm domain in magenta, the fingers domain in green, the DNA primer in gray, the DNA template in yellow, the incoming NTP in cyan, two divalent metal ions in dark blue, and three catalytic carboxylates (D190, D192, D256) in magenta. Secondary structures in pol β palm domain are labeled as in the BstCCA head domain. In the BstCCA structure, the palm-homologous region is shown in magenta and the additional structural features in the head domain are colored red; the protein template region is shown in yellow and the rest of the neck domain is in green; the modeled RNA primer is shown in gray, the incoming CTP in cyan, one divalent metal ion in dark blue, and three carboxylates (D40, D42, E79) in magenta. The RNA primer is modeled from the pol β ternary complex structure based on the superimposition of pol β palm domain and BstCCA head domain. The locations of conserved residues 111–116 are marked with black dots.

(B) Structure-based sequence alignment for BstCCA and pol β in the N-terminal region containing five β strands and three α helices. The structure alignment was done using O (Jones et al., 1991). Structurally equivalent residues in two enzymes are colored red. Catalytically important carboxylate residues are in bold and red. The secondary structures of BstCCA are displayed on top of the alignment.

is a third carboxylate that is directly involved in binding the catalytically essential metal ions needs to be tested by further mutagenesis and structural studies. Two other conserved residues in the NT signature motif, Gly27 and Arg30, bind the triphosphate of ATP or CTP (Figure 3). Both the conformation of the bound CTP and the arrangement of three carboxylates in BstCCA binary complex structure correspond well with those observed in the structure of the active pol β ternary complex (Figure 2A). Despite the appearance of only one divalent metal ion in these structures of the BstCCA, it is likely that BstCCA employs the same two-metal-ion catalytic mechanism as used by pol β and other known polynucleotide polymerases (Brautigam and Steitz, 1998).

The two extra β strands in the head domain of the BstCCA are unique among polymerases and appear to contribute to the binding of the incoming NTP. The conserved residues in the loop region that connects the two extra β strands (6, 7) to the other five β strands play important roles in the enzymatic activity of BstCCA. A highly conserved Arg111 plays an important role in discriminating between ribose and deoxyribose by hydrogen bonding to the 2′ OH group of ATP or CTP (Figure 3). Presumably, the energetic cost of burying the guanidinium group of Arg111 under a deoxyribose without a compensating 2′ OH hydrogen bond makes the binding and incorporation of deoxyribose nucleotides unfavorable. The side chain of another highly conserved residue, Asp112, stacks under the 6 member ring of the ATP base, but not under the CTP base (Figure 3). This purine-specific base stacking provided by Asp112 may assist in discriminating ATP from CTP, explaining why apo-CCA-adding enzymes bind ATP much more tightly than CTP and why the CCA-adding enzyme in *E. coli*

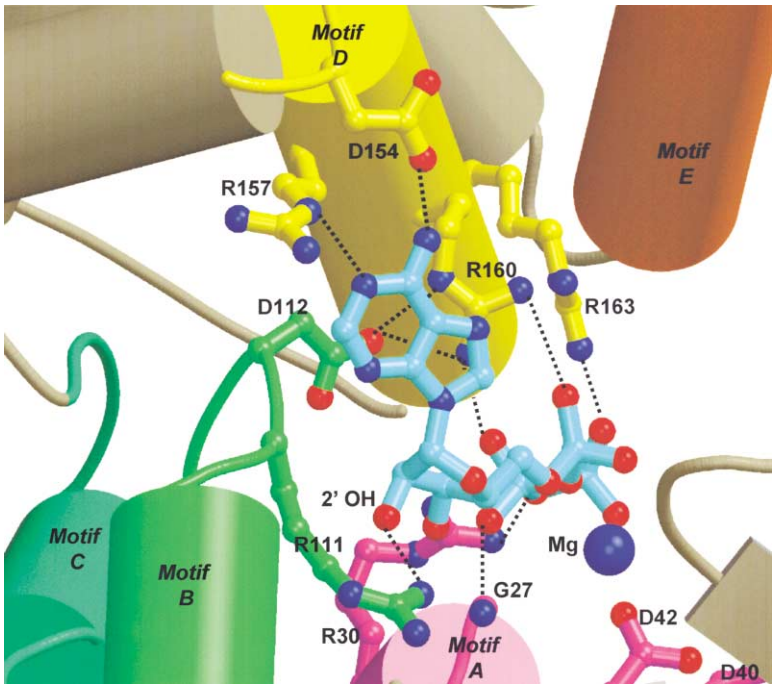

Figure 3. Important Interactions between ATP and Conserved Residues of BstCCA that Confer Specificity for the Base, the Ribose, and the Triphosphate

The five conserved sequence motifs, as explained in detail in Figure 5, are: A (magenta), B (light green), C (dark green), D (yellow), and E (brown).

exists in an ATP bound form (Tomari et al., 2000). Mutation of the residue in the *E. coli* CCA-adding enzyme that is equivalent to Asp112 abolishes A addition but has no effect on C addition (Seth et al., 2002). The purine-specific base stacking between an aspartate and the incoming NTP has not been observed previously in polymerases. Both Arg111 and Asp112 are stabilized by several other conserved residues in the adjacent region (Figure 2A). In sum, the additional structural features in the head domain function to provide discrimination between ribose and deoxyribose as well as purine-specific base stacking.

Protein Template Formed by the Neck Domain

The neck domain of BstCCA provides a template that can specifically recognize either ATP or CTP and discriminate against other nucleotides, ensuring that only ATP and CTP are bound and added to the 3′ terminus of tRNAs. The protein template has three conformations; its conformations in the apo-enzyme and CTP complex are similar. In both ATP and CTP complexes, Asp154 assumes the same conformation and hydrogen bonds either with the 6-NH2 of ATP or the 4-NH2 group of CTP (Figure 4). In the ATP complex, Arg157 hydrogen bonds with the 1-N of ATP while being stabilized by a hydrogen bond with Glu153. In the CTP complex, the side chains of both Arg157 and Glu153 rotate toward CTP, allowing Arg157 to form two hydrogen bonds with both the 3-N and 2-O of CTP. Unlike a nucleic acid template, the protein template can switch its base specificity between ATP and CTP. Unanswered by the present structures, however, is how the tRNA substrate is able to stabilize either the CTP or ATP binding conformation, depending on whether its 3′ terminus is nucleotide 73 (also called the discriminator base), C74, or C75.

Remarkably, the specific patterns of hydrogen donors and acceptors between the protein and ATP or CTP are identical to those seen in the U:A or G:C Watson-Crick base pairing, which readily explains why GTP and UTP are not substrates for CCA-adding enzymes. The Watson-Crick face of either G or U is not complementary to the protein template because their patterns of hydrogen donors and acceptors are the reverse of ATP and CTP. Indeed, when crystals were soaked in solutions containing either GTP or UTP and data to 3 Å resolution were measured, difference electron density was not observed for either, showing that GTP and UTP do not bind to these crystals (data not shown). The protein template can also explain the specificity of NTP analog incorporation by other class II CCA-adding enzymes. For example, the modification of the 6-NH2 of ATP or 4-NH2 of CTP, which dramatically decreases incorporation into tRNAs (Sprinzl et al., 1977), would disrupt the interactions with the protein template seen here.

Unlike a nucleic acid template, the protein template is dynamic. The conformational change in Arg157 defines the size and specificity of the NTP binding pocket. Glu153 plays an important role in nucleotide selection by stabilizing Arg157 in two different conformations. In the *A. aeolicus* CC-adding enzyme, the residue that is equivalent to Glu153 is an arginine. This change may constrain the size of the NTP binding pocket to favor the CTP binding conformation in the CC-adding enzyme; in order for the CC-adding enzyme to accommodate ATP, which has a larger base than CTP, into the NTP binding pocket, the base-pairing arginine would have to rotate away from the NTP binding pocket, which may cause charge repulsion between two arginines. The precise and dynamic protein template, along with ribose recognition and purine-specific base stacking, are the nucleotide selection strategies used by CCA-adding enzymes.

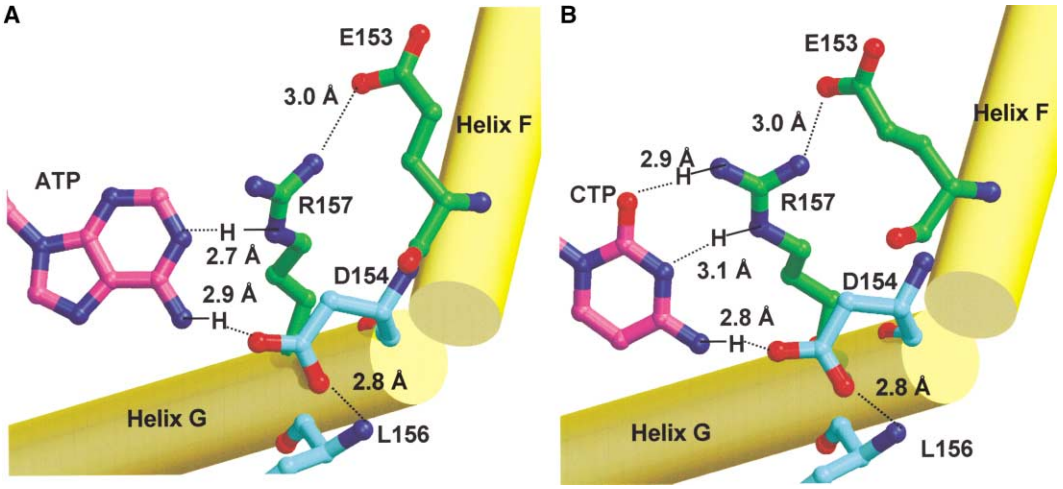

Figure 4. Base-Specific Interactions between BstCCA and ATP or CTP

(A) Interactions between the carboxylate of D154 and the guanidinium group of R157 of BstCCA and ATP.

(B) Interaction between D154 and R157 of BstCCA and CTP. The position of the guanidinium group of R157 is changed to be complementary to C while D154 interacts with the N4 of C in much the same way as it interacts with the N6 of A.

Other Structural Features and Function of the Neck Domain

Although BstCCA neck domain has a templating function that does not exist in any other polymerases, it also resembles a traditional fingers domain in some respects. For example, Arg160 and Arg163 from the neck domain bind and neutralize the triphosphate moiety of ATP or CTP (Figure 3). Arg160 also stabilizes the CTP base or the 5 member ring of the ATP base by hydrophobic interactions that are base nonspecific (Figure 3). Furthermore, Arg160 hydrogen bonds to Asp112 from the head domain, together providing an immobilized stacking platform for ATP or CTP. In T7 DNA polymerase, Tyr527 provides the base-nonspecific stacking platform for the incoming NTP, but the position of Tyr527 is not fixed (Doublie et al., 1998). The movement of Tyr527 in T7 DNA polymerase is believed to be involved in the open-close conformational change of the fingers domain and is an important strategy for pol I enzymes to select for the incoming dNTP that correctly base-pairs with the DNA template. In HIV reverse transcriptase (RT), a mobile Arg72 plays an analogous role (Huang et al., 1998). The rest of the BstCCA neck domain stabilizes the protein template, a function reminiscent of a traditional fingers domain that stabilizes the single-stranded portion of its DNA template.

Structure and Function of the Body and Tail Domains

The body and tail domains at the C terminus extend far from the active site and may be involved in tRNA binding. Footprinting and phosphate modification interference experiments showed that the CCA-adding enzymes interact primarily with the acceptor stem and the T stem loop of tRNA, although a few phosphates of the variable loop of the tRNA also seem critical; moreover, the position of the tRNA on the CCA-adding enzymes does not change during CCA addition (Shi et al., 1998). CCA-adding enzymes must recognize structural features common to tRNAs. In BstCCA structures, the body and tail domains are about 60 Å long, which is similar to the length of protein-protected region of tRNAs and supports a role of these domains in tRNA binding. Although the tail domain is distant from the active site and flexibly attached to the rest of the protein structure, it plays an important role in the CCA addition because deletion of the last two corresponding α helices in the *E. coli* CCA-adding enzyme abolishes both A addition and C addition (Zhu and Deutscher, 1987). We hypothesize that the acceptor stem and the T stem loop of tRNA are bound in parallel with the extended C-terminal domains of BstCCA, with the body domain binding the acceptor and T stems of tRNA and the tail domain binding the T loop of tRNA. This model is consistent with footprinting and phosphate modification interference data, molecular shapes of BstCCA and tRNA, and truncation experiments done on the *E. coli* CCA-adding enzyme.

Conserved Motifs in Eubacterial CCA-Adding Enzymes and Other Class II Enzymes

The head and neck domains of the BstCCA structure contain five motifs (motifs A–E in Figures 5 and 3) that are highly conserved in the class II NT enzymes, including eubacterial and eukaryotic CCA-adding enzymes, *A. aeolicus* CC-adding enzyme and A-adding enzyme and eubacterial PAPs, indicating that these two domains are structurally conserved for class II enzymes. Motif A is the signature motif for all members of the NT superfamily and includes the two carboxylates that bind the catalytically essential metal ions as well as a helix-turn structure that binds the triphosphate moiety of the incoming NTP. The other four motifs are unique to the class II enzymes. Motif B functions in ribose recognition and purine-specific base stacking with the incoming NTP. Motif C connects the head and neck domains. Motif D functions as the template for the incoming NTP, provides a base-nonspecific stacking platform for the incoming NTP and

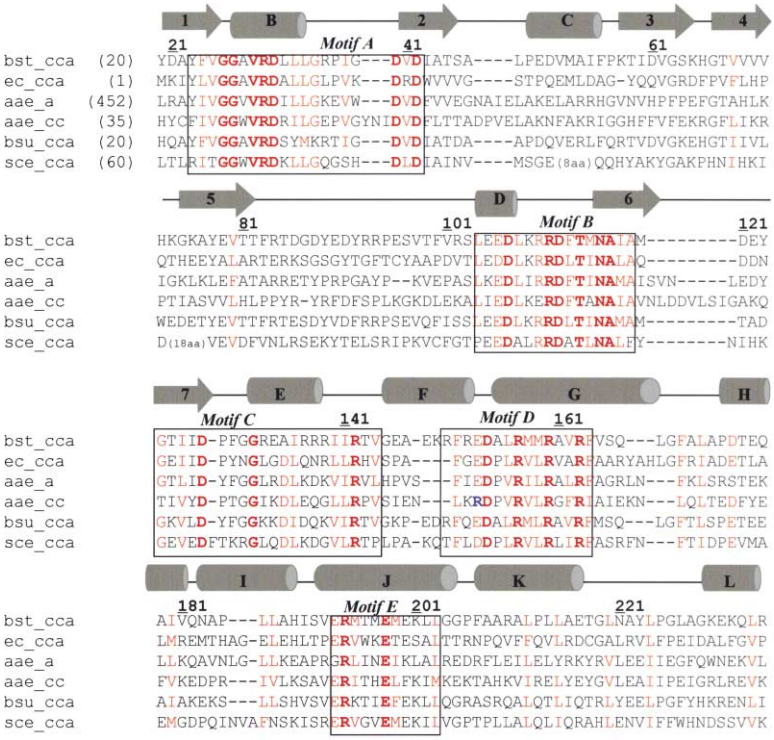

Figure 5. Multiple Sequence Alignment in the N-Terminal Region for Class II Enzymes Including Eubacterial and Eukaryotic CCA-Adding Enzymes, *A. aeolicus* CC-Adding Enzyme and A-Adding Enzyme, and Eubacterial PAPs

The initial alignment was generated by the ClusterW program and then adjusted manually. Only absolutely conserved regions (in bold and red) and highly conserved regions (in red) are shaded. Five highly conserved motifs are boxed. In motif C of *Aquifex aeolicus* CC-adding enzyme, the arginine residue that may be responsible for the absence of ATP binding conformation is shown in blue. Abbreviations: bst_cca for *Bacillus stearothermophilus* CCA-adding enzyme, ec_cca for *Escherichia coli* CCA-adding enzyme, aae_a for *Aquifex aeolicus* A-adding enzyme, aae_cc for *Aquifex aeolicus* CC-adding enzyme, bsu_cca for *Bacillus subtilis* CCA-adding enzyme, and sce_cca for *Saccharomyces cerevisiae* CCA-adding enzyme. The secondary structures of BstCCA are displayed on top of the alignment.

binds the triphosphate moiety. Motif E stabilizes the helix-turn structure in motif D and may also be involved in interacting with the RNA primer strand (described below). Those residues that are directly involved in the interactions with ATP or CTP in BstCCA structures are highly conserved among all class II enzymes, indicating that all class II enzymes employ similar nucleotide selection strategies for their template-independent polymerization activities.

Reevaluation of Existing Models on the Mechanism of Sequential CCA Addition

Although the structures of ternary complexes of the CCA-adding enzyme with tRNA substrates and NTP will

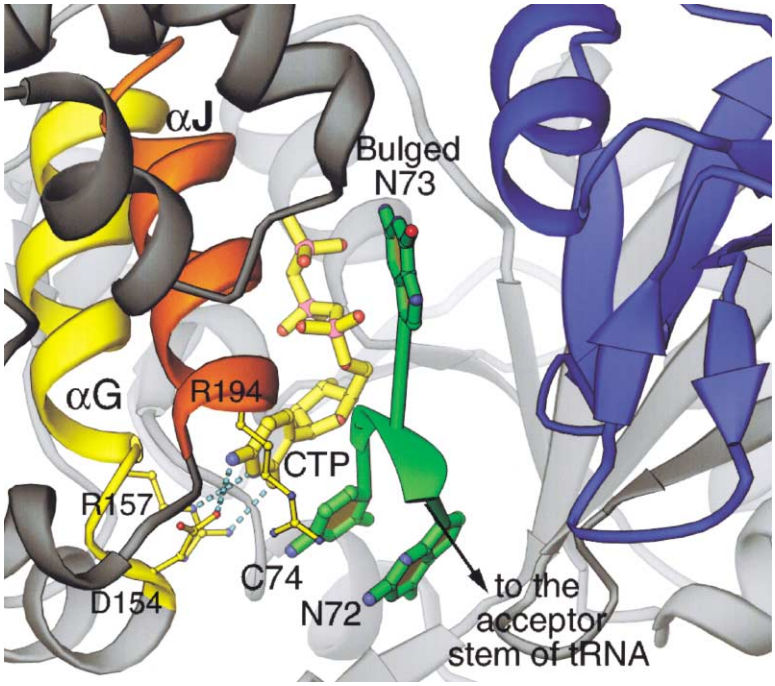

Figure 6. A Model Showing a Singly Scrunched tRNA Primer Displaying a Single Bulged-Out Nucleotide

The model was created by superimposing the 3′ A76 of the terminal four nucleotide of $tRNA^{Gln}$, as observed in complex with its cognate synthetase (Rould et al., 1989), on the bound ATP and then changing the bases to CTP, C74, G73, and G72 (C74 and N72 are also positioned identically to the primer terminus homology modeled from pol β shown in Figure 2). This modeling positions the backbone phosphate of the RNA (green ribbon) near a highly conserved Arg194 (yellow) on a highly conserved Helix J (orange). Helix J is adjacent to Helix G (yellow), whose Asp154 and Arg157 make base-specific interactions with the incoming CTP (yellow) that will become C75. One subunit of the dimeric enzyme is shown in gray and the other in blue.

be required to define the detailed mechanism for specific CCA addition, the present structures impose constraints that eliminate or modify previous models while suggesting additional possibilities. Models that postulated multiple ATP and CTP binding sites aligned in a row (e.g., Deutscher, 1972, 1982; Tomari et al., 2000; Seth et al., 2002) can be rejected because ATP and CTP bind to the same pocket in our structures, consistent with earlier observations that mutation of the catalytic carboxylates eliminated both ATP and CTP incorporation (Yue et al., 1998). A second, "collaborative templating" model postulates a progressive refolding of the growing 3′ end of the tRNA that becomes directly involved in templating the next incoming nucleotide (Yue et al., 1998; Shi et al., 1998). While the 3′ end is very likely to refold as elongation proceeds, the structures presented here show that the protein itself is sufficient to provide a template specific for ATP or CTP without direct "collaboration" with the tRNA. The third, "scrunching shuttling" model, likewise needs to be modified. This model postulated that the 3′ end refolded and scrunched upon addition of C74 but that after the addition of C75, the 3′ terminus might be long enough to "shuttle" to the active site of a second monomer whose structure was specific for addition of ATP. The shuttling portion of this model can be eliminated because the two active sites observed in this crystalline dimer are too far separated for a 3′ C75 to reach the second active site without significant melting of the tRNA (at least three base pairs). Indeed, there is a "ridge" of protein structure between the two dyad related active sites so that the surface distance between them is more than 40 Å. A fully extended, unstacked terminal three nucleotides (73, 74, and 75) could only reach about 21 Å.

A Modified Scrunching Model for Sequential CCA Addition

It remains possible, however, that a single active site can incorporate two C's with the protein template in the CTP-adding conformation and the terminal A76 with the protein template in the ATP-adding conformation and that the growing 3′ terminus is "scrunched" to allow identical positioning of the 3′ terminus after each addition. The "scrunching" would be analogous to that observed for the template strand of a transcribing RNA polymerase bound to an immobile promoter DNA during the initiation phase of transcription, where two nucleotides of template are scrunched into an enzyme pocket after formation of a trinucleotide transcript (Cheetham and Steitz, 1999). Such a model for CCA-adding activity would require that the active site be able to accommodate a compactly refolded 3′ terminus that is correctly positioned for addition of the next nucleotide. The position of nucleotide 72 and the rest of the tRNA remain unmoved during nucleotide incorporation, as suggested by modification interference and footprinting experiments (Shi et al., 1998). The single-stranded 3′ end of the tRNA scrunches twice before the scrunching pocket for the enzyme is full and the scrunched primer presumably somehow triggers a conformational change that switches the specificity of the protein template from CTP to ATP (Figure 6).

To understand the length specificity of polymerization by CCA-adding enzymes as well as the viability of a scrunching hypothesis, we examined the size of the pocket that could accommodate a scrunched RNA. The potential scrunching pocket is formed by the highly conserved Helix J and the RNA primer strand that can be positioned by homology modeling from the pol β ternary complex structure (Figure 2A). The scrunched nucleotides would flip out of the elongating primer strand and can fit into a pocket (Figure 6). Such scrunched or "bulged" RNA structures are not uncommon, since they are seen in many existing RNA crystal structures such as those of the TAR RNA (Ippolito and Steitz, 1998), the hepatitis delta virus ribozyme (Ferre-D'Amare et al., 1998), the ribosome large subunit (Ban et al., 2000), and the 3′ tRNA terminus in the $tRNA^{Gln}$ when complexed with its cognate synthetase (Rould et al., 1989). Our efforts to model these bulged RNA structures into the potential scrunching pocket indicate that both singly and doubly bulged RNA primer termini-modeled models could fit well into the pocket, whereas the model of a triply scrunched RNA primer clashes with Helix J. The limited size of the potential scrunching pocket provides the possibility that CCA-adding enzymes would sterically preclude the scrunching of a third nucleotide, thereby preventing the extension beyond CCA. Consistent with the modeling, kinetic studies show that binding of a second ATP to the active site of CCA-adding enzymes kicks out the tRNA with a completed CCA end (Evans and Deutscher, 1972; Miller and Philipps, 1973).

The mechanism by which the enzyme detects the length of the 3′ terminus and switches the templating specificity from C to A after addition of C75 is a matter of speculation at this point. However, we note that the highly conserved Arg194 from motif E lies in Helix J and is positioned near the backbone of the primer terminus (Figure 6). Helix J, which is part of the potential scrunching pocket, is adjacent to Helix G, which is part of the protein template, so that changes in the state of the primer strand could in some manner be conveyed through Helix J to Helix G to alter the templating specificity. Clearly, the exact mechanism can only be derived from the crystal structures of appropriate complexes.

Experimental Procedures

Protein Purification and Crystallization

The gene for the *Bacillus stearothermophilus* CCA-adding enzyme (BstCCA) was cloned by degenerate PCR and joined to a C-terminal histidine tag (H.D.C., K. Tomita, and A.M.W., submitted). The histidine-tagged enzyme was overexpressed and purified essentially as described for the *S. shibatae* CCA-adding enzyme (Li et al., 2000), except that a Ni-NTA column was added to the purification procedure after the heat treatment step. Selenomethionine was incorporated into the protein using an auxotrophic strain of *E. coli* and M9 growth media in which selenomethione replaced methionine. Crystals grew spontaneously when the protein at 8 mg/ml concentration was dialyzed in a buffer containing 50 mM Tris (pH 7.5), 200 mM NaCl, and 10 mM $MgCl_2$. Crystals were transferred into a cryoprotectant solution containing 20% glycerol before flash cooling in liquid propane. ATP or CTP was soaked into crystals at 2 mM concentration each.

Data Collection, Structure Determination, and Refinement

MAD and native X-ray diffraction data were collected at the Advanced Photon Source (Argonne, IL) beamline 14-ID and the NSLS beamline X25 (Brookhaven, NY) and processed using the HKL pack-

age (Otwinowski and Minor, 1997). The crystals contain two protein molecules per asymmetric unit, giving a crystal solvent content of 61%. Phasing was done using the CCP4 package (CCP4, 1994). An ethyl mercuric phosphate derivative was used to generate single isomorphous replacement phases that were used to calculate an anomalous difference Fourier map in which all 20 selenomethionine sites in the selenomethinine-labeled crystal were located. MAD phasing calculated using selenomethionine data generated interpretable electron density maps. Solvent flattening and noncrystallographic symmetry (NCS) averaging significantly helped phase calculation. The data were also sharpened by applying a negative B factor (−70) operation using the cad program in CCP4 to scale up the contributions of high-resolution reflections. This procedure improved the quality of electron density maps. Model building was done using O (Jones et al., 1991). Structural refinement was done using both CNS (Brunger et al., 1998) and CCP4. The final refined models contain two protein molecules, each including 395 of the 404 protein residues (excluding nine residues from 88 to 96 on a disordered loop). One divalent metal ion was located from the $F_o -$ F_o difference Fourier map between a crystal soaked in 10 mM $MnCl_2$ and a crystal of the apo-protein.

Acknowledgments

We thank Satwik Kamtekar for helpful discussions. We thank beamline 14-ID at the Advanced Photon Source and beamline X25 at the NSLS for help in data collection. This work was supported by NIH grants GM57510 (to T.A.S.) and GM59804 (to A.M.W.).

Received: August 28, 2002
Revised: October 2, 2002

References

Aebi, M., Kirchner, G., Chen, J.Y., Vijayraghavan, U., Jacobson, A., Martin, N.C., and Abelson, J. (1990). Isolation of a temperature-sensitive mutant with an altered tRNA nucleotidyltransferase and cloning of the gene encoding tRNA nucleotidyltransferase in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. *265*, 16216–16220.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science *289*, 905–920.

Bard, J., Zhelkovsky, A.M., Helmling, S., Earnest, T.N., Moore, C.L., and Bohm, A. (2000). Structure of yeast poly(A) polymerase alone and in complex with 3′-dATP. Science *289*, 1346–1349.

Brautigam, C.A., and Steitz, T.A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. Curr. Opin. Struct. Biol. *8*, 54–63.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

Cheetham, G.M., and Steitz, T.A. (1999). Strauctures of a transcribing T7 RNA polymerase initiation complex. Science *286*, 2305–2309.

Chothia, C., and Janin, J. (1975). Principles of protein-protein recognition. Nature *256*, 705–708.

Delarue, M., Boule, J.B., Lescar, J., Expert-Bezancon, N., Jourdan, N., Sukumar, N., Rougeon, F., and Papanicolaou, C. (2002). Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. EMBO J. *21*, 427–439.

Deutscher, M.P. (1972). Reactions at the 3′ terminus of transfer ribonucleic acid. 3. Catalytic properties of two purified rabbit liver transfer ribonucleic acid nucleotidyl transferases. J. Biol. Chem. *247*, 459–468.

Deutscher, M.P. (1982). tRNA nucleotidyltransferase. The Enzymes *15*, 183–215.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature *391*, 251–258.

Evans, J.A., and Deutscher, M.P. (1978). Kinetic analysis of rabbit liver tRNA nucleotidyltransferase. J. Biol. Chem. *253*, 7276–7281.

Ferre-D'Amare, A.R., Zhou, K., and Doudna, J.A. (1998). Crystal structure of a hepatitis delta virus ribozyme. Nature *395*, 567–574.

Green, R., and Noller, H.F. (1997). Ribosomes and translation. Annu. Rev. Biochem. *66*, 679–716.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science *282*, 1669–1675.

Holm, L., and Sander, C. (1995). DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. Trends Biochem. Sci. *20*, 345–347.

Holm, L., and Sander, C. (1998). Touring protein fold space with Dali/FSSP. Nucleic Acids Res. *26*, 316–319.

Ippolito, J.A., and Steitz, T.A. (1998). A 1.3-Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. Proc. Natl. Acad. Sci. USA *95*, 9819–9824.

Jones, T.A., Zou, J.Y., and Cowan, S.W. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Li, F., Wang, J., and Steitz, T.A. (2000). *Sulfolobus shibatae* CCA-adding enzyme forms a tetramer upon binding two tRNA molecules: a scrunching-shuttling model of CCA specificity. J. Mol. Biol. *304*, 483–492.

Martin, G., and Keller, W. (1996). Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. EMBO J. *15*, 2593–2603.

Martin, G., Keller, W., and Doublie, S. (2000). Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. EMBO J. *19*, 4193–4203.

Miller, J.P., and Philipps, G.R. (1973). Transfer ribonucleic acid nucleotidyltransferase from *Escherichia coli*. 3. Kinetic analysis. J. Biol. Chem. *246*, 1280–1284.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920–930.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature *313*, 762–766.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science *264*, 1891–1903.

Rould, M.A., Perona, J.J., Soll, D., and Steitz, T.A. (1989). Structure of *E. coli* glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. Science *246*, 1135–1142.

Sakon, J., Liao, H.H., Kanikula, A.M., Benning, M.M., Rayment, I., and Holden, H.M. (1993). Molecular structure of kanamycin nucleotidyltransferase determined to 3.0 Å resolution. Biochemistry *32*, 11977–11984.

Seth, M., Thurlow, D.L., and Hou, Y.M. (2002). Poly(C) synthesis by class I and class II CCA-adding enzymes. Biochemistry *41*, 4521–4532.

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

Sprinzl, M., and Cramer, F. (1979). The -C-C-A end of tRNA and its role in protein biosynthesis. Prog. Nucleic Acid Res. Mol. Biol. *22*, 1–16.

Sprinzl, M., Sternbach, H., von der Haar, F., and Cramer, F. (1977). Enzymatic incorporation of ATP and CTP analogues into the 3′ end of tRNA. Eur. J. Biochem. *81*, 579–589.

Tomari, Y., Suzuki, T., Watanabe, K., and Ueda, T. (2000). The role of tightly bound ATP in *Escherichia coli* tRNA nucleotidyltransferase. Genes Cells *5*, 689–698.

Tomita, K., and Weiner, A.M. (2001). Collaboration between CC- and A-adding enzymes to build and repair the 3′-terminal CCA of tRNA in *Aquifex aeolicus*. Science *294*, 1334–1336.

Yue, D., Weiner, A.M., and Maizels, N. (1998). The CCA-adding enzyme has a single active site. J. Biol. Chem. *273*, 29693–29700.

Zhu, L., and Deutscher, M.P. (1987). tRNA nucleotidyltransferase is not essential for *Escherichia coli* viability. EMBO J. *8*, 2473–2477.

Accession Numbers

Coordinates of the apo-protein and its complexes with ATP and CTP as well as the experimental diffraction amplitudes have been deposited in the Protein Data Bank with accession numbers 1MIW (ATP complex), 1MIY (CTP complex), and 1MIV (apo-protein).