Substrate Affinity Versus Catalytic Efficiency: Ancestral Sequence Reconstruction of tRNA Nucleotidyltransferases Solves an Enzyme Puzzle

Martina Hager,[†,1] Marie-Theres Pöhler,[†,1] Franziska Reinhardt,[2,3] Karolin Wellner,[1] Jessica Hübner,[2] Heike Betat,[1] Sonja Prohaska,[2,3,4,5] and Mario Mörl [*,1]

[1]Institute for Biochemistry, Leipzig University, Brüderstraße 34, D-04103 Leipzig, Germany

[2]Computational EvoDevo Group, Institute for Computer Science, Leipzig University, Härtelstr. 16-18, 04109 Leipzig, Germany

[3]Interdisciplinary Centre for Bioinformatics, Leipzig University, Härtelstr. 16-18, 04109 Leipzig, Germany

[4]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

[5]Complexity Science Hub Vienna, Josefstädter Str. 39, 1080 Wien, Austria

[†]These authors contributed equally to this work.

***Corresponding author:** E-mail: mario.moerl@uni-leipzig.de.

Associate editor: Li Liu

Abstract

In tRNA maturation, CCA-addition by tRNA nucleotidyltransferase is a unique and highly accurate reaction. While the mechanism of nucleotide selection and polymerization is well understood, it remains a mystery why bacterial and eukaryotic enzymes exhibit an unexpected and surprisingly low tRNA substrate affinity while they efficiently catalyze the CCA-addition. To get insights into the evolution of this high-fidelity RNA synthesis, the reconstruction and characterization of ancestral enzymes is a versatile tool. Here, we investigate a reconstructed candidate of a 2 billion years old CCA-adding enzyme from Gammaproteobacteria and compare it to the corresponding modern enzyme of *Escherichia coli*. We show that the ancestral candidate catalyzes an error-free CCA-addition, but has a much higher tRNA affinity compared with the extant enzyme. The consequence of this increased substrate binding is an enhanced reverse reaction, where the enzyme removes the CCA end from the mature tRNA. As a result, the ancestral candidate exhibits a lower catalytic efficiency in vitro as well as in vivo. Furthermore, the efficient tRNA interaction leads to a processive polymerization, while the extant enzyme catalyzes nucleotide addition in a distributive way. Thus, the modern enzymes increased their polymerization efficiency by lowering the binding affinity to tRNA, so that CCA synthesis is efficiently promoted due to a reduced reverse reaction. Hence, the puzzling and at a first glance contradicting and detrimental weak substrate interaction represents a distinct activity enhancement in the evolution of CCA-adding enzymes.

Key words: **tRNA nucleotidyltransferase, CCA-adding enzyme, ancestral sequence reconstruction, RNA polymerization, RNA interaction.**

Article

Introduction

Ancestral sequence reconstruction (ASR) is a valuable tool to infer the evolution of specific features of a modern protein. Originally described as a molecular approach by Pauling and Zuckerkandl (Pauling et al. 1963), Benner and coworkers used ASR in a pioneering work to reconstruct an ancestral ribonuclease (Stackhouse et al. 1990; Jermann et al. 1995). Since then, many predecessors of modern enzymes were generated and biochemically investigated, and although these enzymes represent only approximations toward ancestral enzymes, they are very useful tools revealing important information on functionality, specificity, velocity, and other parameters of their extant counterparts (Peltier et al. 2000; Thomson et al. 2005; Carrigan et al. 2015; Risso et al. 2015; Merkl and Sterner 2016b; Sánchez-Ruiz and Risso 2018). Here, we investigated a candidate for an ancestral tRNA nucleotidyltransferase from Gammaproteobacteria with a corresponding age of about 2 billion years, when this class of bacteria evolved (Battistuzzi et al. 2004; Williams et al. 2010).

tRNA nucleotidyltransferases (CCA-adding enzymes) are ubiquitous and highly specialized RNA polymerases that add the invariant CCA terminus to the tRNA 3′-end in all kingdoms of life (Xiong and Steitz 2006; Betat et al. 2010; Vörtler and Mörl 2010). Polymerization of this sequence is a highly accurate process and is discussed to originate from a primordial telomerase function in the RNA world (Maizels and Weiner 1994, 1999; Maizels et al. 1999; Weiner and Maizels 1999). According to sequence and structural composition, the responsible enzymes are

divided into two different groups. While archaea carry class 1 enzymes, eukaryotes, and bacteria possess class 2. In the latter, a set of templating amino acid residues in a single nucleotide binding pocket form Watson–Crick-like hydrogen bonds to the incoming NTP (Li et al. 2002). After the addition of two C residues to the tRNA, a reorganization of this amino acid template dictates a specificity switch from C toward A incorporation (Li et al. 2002; Neuenfeldt et al. 2008; Toh et al. 2009; Betat et al. 2010; Tomita and Yamashita 2014).

While class 2 enzymes catalyze the CCA polymerization at high fidelity and efficiency, it is a mystery why these enzymes exhibit a rather weak affinity toward the tRNA substrate, so that in most cases, binding constants could not be determined (Shi, Maizels, et al. 1998; Tretbar et al. 2011; Ernst et al. 2015; Hennig et al. 2020). In contrast, in class 1 enzymes, binding constants in a range of 30–700 nM are described (Shi, Maizels, et al. 1998; Okabe et al. 2003; Cho et al. 2005, 2008). To investigate the molecular reason for these obviously conflicting features of catalytic efficiency versus low substrate affinity, we set out to compare class 2 enzymes that differ in these parameters. However, only enzymes adapted to bizarre armless tRNA structures (Hennig et al. 2020) or unconventional enzymes with a series of Q/N insertions of unknown function exhibit a detectable affinity toward tRNA substrates (Erber et al. 2020). Hence, we used the approach of ASR to generate a conventional candidate enzyme with ancestral features that differ from modern enzymes in terms of substrate binding, while it is not adapted to bizarre tRNA substrates.

For this reconstruction, we focused on tRNA nucleotidyltransferase sequences from Gammaproteobacteria, as the *Escherichia coli* enzyme (*Eco*CCA) is one of the best characterized enzymes of this class (Cudny et al. 1986; Zhu and Deutscher 1987; Shi, Maizels, et al. 1998; Shi, Weiner, et al. 1998; Hou 2000; Betat et al. 2004; Kim et al. 2009; Hoffmeier et al. 2010). Furthermore, only one gammaproteobacterial species is known carrying more than a single tRNA nucleotidyltransferase, so that false annotations can be avoided (Tretbar et al. 2011). In addition, Gammaproteobacteria exhibit a robust phylogeny, facilitating ASR (Williams et al. 2010).

In our analysis, the selected candidate enzyme turned out to be active and showed remarkable properties. While fidelity and specificity of the ancestral enzyme are comparable to its modern counterparts, its binding to tRNA substrates is unexpectedly high and differs dramatically from modern gammaproteobacterial tRNA nucleotidyltransferases. This tight interaction results in an enhanced reverse reaction and an equally strong binding to substrate (tRNA) as well as product (tRNA-CCA). As a consequence, the ancestral enzyme candidate exhibits a lower turnover rate in vitro and a reduced functionality in vivo. Hence, our results demonstrate that the extant class 2 CCA-adding enzymes optimized their catalytic performance by reducing the tRNA substrate affinity, so that substrate inhibition as well as reverse reaction are minimized and allow for an efficient polymerization (forward) reaction.

Results

Sequence Reconstruction of Ancestral CCA-Adding Enzyme Candidates

ASR is a four-step process that requires careful evaluation of intermediate results, including taxon sampling, multiple sequence alignment, phylogenetic tree reconstruction, and ancestor reconstruction (Merkl and Sterner 2016a). Based on experimentally verified bacterial class 2 CCA-adding enzymes, homologous protein sequences were identified from representative gammaproteobacterial RefSeq genomes (O'Leary et al. 2016). Sequence selection aimed at covering as many orders and families as possible, resulting in coverage of 13 of 22 orders and 32 of 66 families from NCBI taxonomy. In order to avoid taxon sampling biases, the number of sequences in the corresponding orders was adjusted manually, resulting in a final set of 102 candidate sequences from 11 Alteromonadales, 7 Cellvibrionales, 11 Chromatiales, 12 Enterobacteriales, 2 Kangiellales, 4 Legionellales, 4 Methylococcales, 10 Oceanospirillales, 5 Pasteurellales, 13 Pseudomonadales, 5 Thiotrichales, 11 Vibrionales, and 7 Xanthomonadales (supplementary table S1, Supplementary Material online). Alignment was carried out using Clustal Omega (Sievers and Higgins 2018) adding four sequences from Betaproteobacteria as an outgroup. The reliability of the alignment was verified using the *Consistency of the Overall Residue Evaluation* implemented in CORE (Notredame et al. 2000; Wallace et al. 2006).

The phylogenetic gene tree required for ASR was computed from the GBlocks-trimmed alignment to improve the phylogenetic signal (fig. 1). The resulting tree is in good agreement with the phylogeny of Gammaproteobacteria reported by Williams et al. (2010). In most cases, sequences from the same order clustered in subtrees with high bootstrap values (95–100). Oceanospirillales and Alteromonadales were found in multiple positions of the tree, which is in agreement with these taxonomic orders not being monophyletic. The deepest splits of the tree had very low bootstrap support with values ranging from 9 to 30. This, however, is expected given the divergence times of gammaproteobacterial orders some 0.8–2 billion years ago (Kumar et al. 2017).

Ancestral sequences were reconstructed using FastML (Ashkenazy et al. 2012). FastML offers two different methods for character reconstruction, that is, joint and marginal reconstruction, and a separate algorithm for indel reconstruction. The results of the two methods for character reconstruction are highly similar (≥90% sequence identity). Indel reconstruction was performed using cutoff values between 0.1 and 1.0 in steps of 0.1. This results in two indels of 4–6 aa at positions 340 and 350. An additional indel at position 270 is reconstructed at cutoff values 0.7 or lower (supplementary fig. S2, Supplementary Material online).

The selection of the best ancestral candidate sequences for biochemical characterization was made using these two indel variants of the joint character reconstruction with indel cutoff 0.7 (*Anc*CCA1) and 0.9 (*Anc*CCA3) and without indel reconstruction (*Anc*CCA2). In addition,

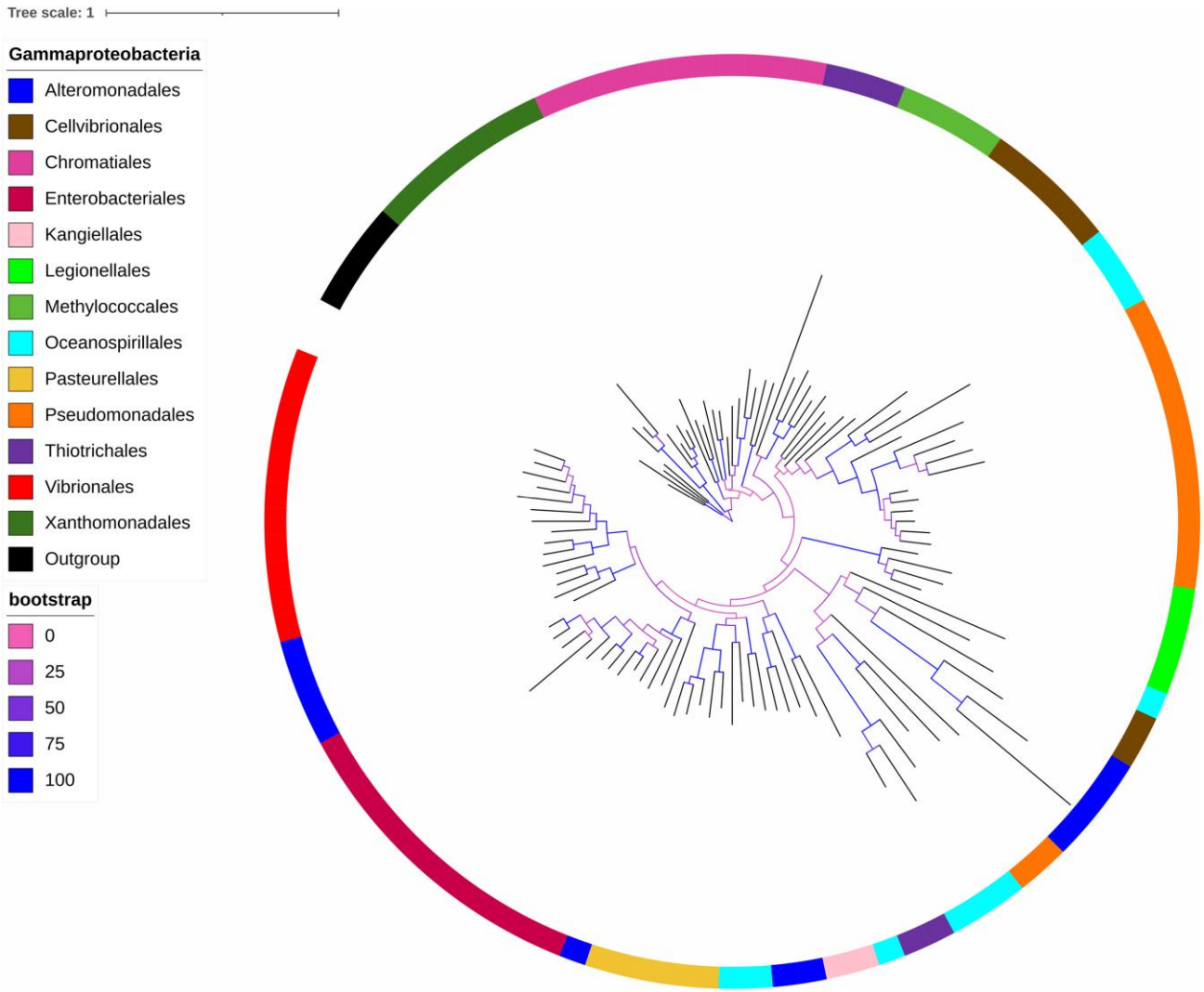

FIG. 1. Phylogenetic tree for ancestral sequence reconstruction of gammaproteobacterial CCA-adding enzymes. The tree is rooted using four betaproteobacterial sequences from putative tRNA nucleotidyltransferases as an outgroup. To each of the 13 orders, an individual color is assigned to locate members of the respective order in the tree. Colors of inner branches indicate bootstrap values. Species names and accession numbers for the respective enzyme sequences are listed in supplementary table S1, Supplementary Material online. A high resolution representation of this figure including species names is shown in supplementary figure S1, Supplementary Material online.

one indel variant of the marginal reconstruction with indel cutoff 0.9 (*Anc*CCA4) was selected (supplementary figs. S2 and S3, Supplementary Material online). All reconstructed candidate sequences carry the conserved catalytic core motifs of class 2 tRNA nucleotidyltransferases (supplementary fig. S3A, Supplementary Material online).

CCA-Adding Activity of Ancestral Enzyme Candidates

The codon-optimized open reading frames of the four reconstructed enzyme candidates were ordered at GenScript (Rijswijk, Netherlands), cloned and recombinantly expressed in *E. coli* BL21 (DE3) Δ*cca*. All recombinant enzymes showed a comparable CCA-adding activity on a radioactively labeled in vitro transcribed standard substrate yeast $tRNA^{Phe}$ (supplementary fig. S3B, Supplementary Material online). As candidate *Anc*CCA2 shows a bipartite C-terminal insertion of 10 amino acids (Positions 317–320 and 326–331; supplementary fig. S3A, Supplementary Material online), this candidate was excluded from further analysis. While candidates *Anc*CCA1, *Anc*CCA3, and *Anc*CCA4 are highly similar in sequence, *Anc*CCA1 showed the highest recombinant expression levels and efficient purification. Hence, it was selected for a detailed characterization. For comparison, the corresponding enzyme of *E. coli* (*Eco*CCA), one of the best characterized class 2 enzymes (Cudny and Deutscher 1986; Zhu and Deutscher 1987; Shi, Maizels, et al. 1998; Shi, Weiner, et al. 1998; Hou 2000; Betat et al. 2004; Kim et al. 2009; Hoffmeier et al. 2010), was chosen as an extant gammaproteobacterial representative. When individual NTPs were offered, both *Anc*CCA1 and *Eco*CCA selectively incorporated two C residues, while the enzymes added a complete CCA end in the presence of all four NTPs (fig. 2A, upper panel). Sequence analysis of the cloned reaction product of *Anc*CCA1 exclusively

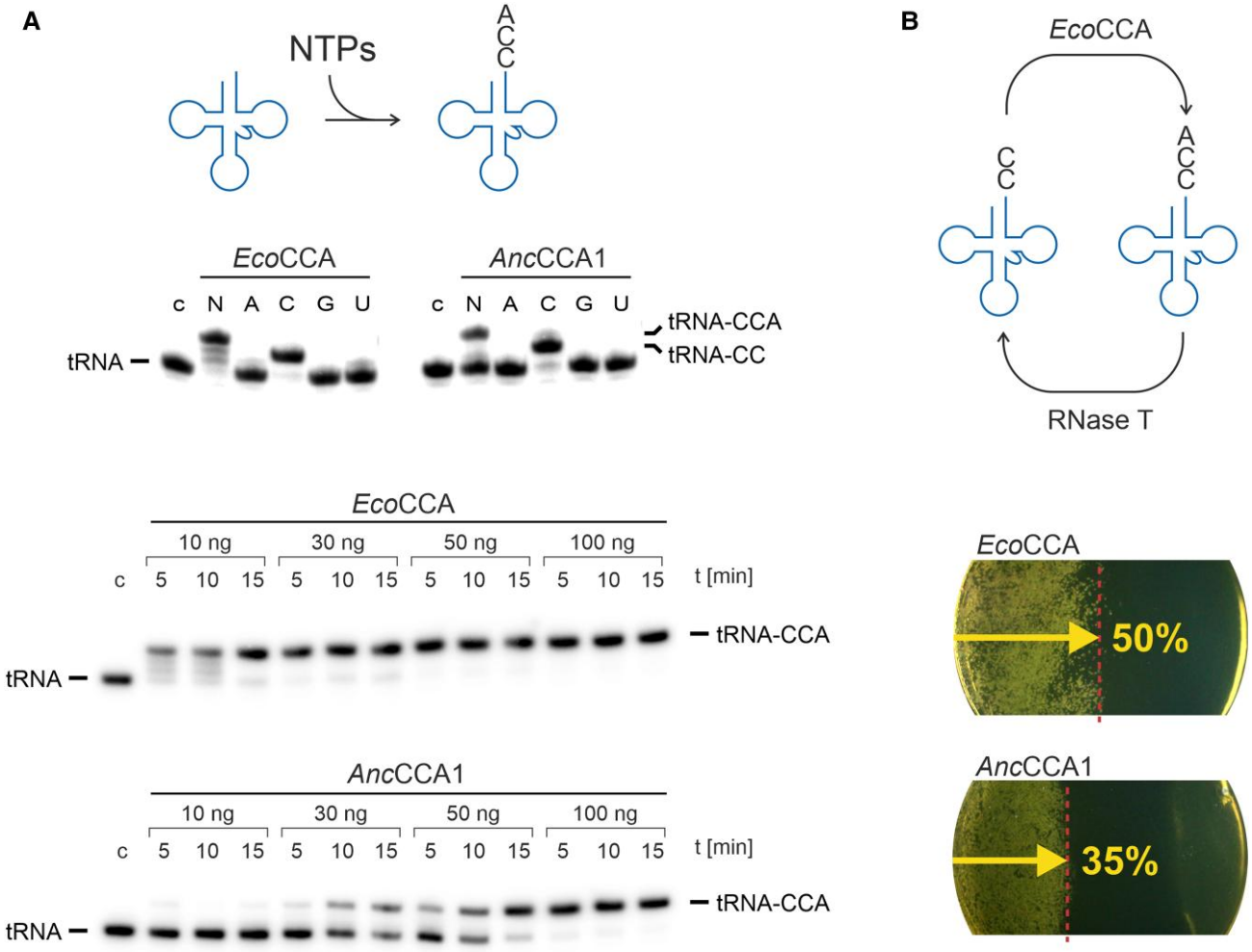

Fig. 2. CCA-addition of *Anc*CCA1 in comparison with *Eco*CCA. (*A*) In vitro activity. Upper panel: Both *Eco*CCA and *Anc*CCA1 selectively add two C residues on a tRNA without CCA end, while other NTPs are not accepted. The full-length product generated by *Anc*CCA1 in the presence of all 4 NTPs (NTP) was isolated and cloned. Sequence analysis of 40 individual clones revealed exclusively correct full or partial CCA ends, indicating that this enzyme has a high fidelity comparable to extant enzymes. Lower panel: CCA-addition with increasing amounts of recombinant enzymes in a time series. Here, *Eco*CCA exhibits a much higher efficiency in polymerization compared with *Anc*CCA1, where products with complete CCA ends occur only at higher enzyme concentrations (30 ng). (*B*) In vivo activity. *E. coli* cells express *Eco*CCA or *Anc*CCA1 in the presence of increasing amounts of RNase T in a linear gradient across the agar plate. While *Eco*CCA expression leads to cell growth of up to 50% of the RNase T gradient, *Anc*CCA1 allows only for growth of up to 35%, indicating that the ancestral candidate is also in vivo less efficient in polymerization. The shown agar plates are representative examples of this analysis. c: Size controls representing tRNA transcripts with or without CCA ends.

showed correct full (27 sequences, 68%) or partial CCA sequences (13 sequences, 32%), indicating that the ancestral candidate exhibits a fidelity comparable to the modern enzyme. Concerning (t)RNA substrates, *Anc*CCA1 shows a comparably high specificity, as it accepts only conventional tRNA structures, but not bizarre substrates such as armless tRNAs recently identified in Enoplea (supplementary fig. S4, Supplementary Material online; Hennig et al. 2020). In addition, substrate RNAs artificially selected for modern CCA-adding enzymes (Wende et al. 2015) were also not accepted by the ancestral candidate. Hence, *Anc*CCA1 is surprisingly less promiscuous in RNA substrate selection compared with its extant counterparts.

To compare the catalytic efficiency of *Eco*CCA and *Anc*CCA1, increasing amounts of recombinant enzymes were incubated with $tRNA^{Phe}$ and NTPs for increasing time (fig. 2A, lower panel). While the *E. coli* enzyme showed a complete CCA-addition already at low concentrations (10 ng) and 5–10 min of incubation, the ancestral candidate was much less active and synthesized CCA ends only at concentrations of 30 ng (10 min) and higher. These differences in catalytic efficiency are in very good agreement with the determined turn-over numbers obtained in steady-state kinetic analyzes, where *Eco*CCA is more efficient than *Anc*CCA1 in both CC- (k_{cat} of $0.82 \pm 0.13\ s^{-1}$ for *Eco*CCA vs. $0.34 \pm 0.04\ s^{-1}$ for *Anc*CCA1) as well as A-addition (k_{cat} of $3.11 \pm 0.34\ s^{-1}$ vs. $0.85 \pm 0.12\ s^{-1}$). However, as such in vitro data frequently do not match in vivo activities, a recently established approach to monitor tRNA A-adding activity in *E. coli* was used, where the expression of RNase T, an enzyme involved in tRNA end turnover, leads to tRNAs with incomplete CCA ends in the cell (Wellner et al. 2019). The catalytic efficiency of a recombinantly expressed CCA-adding enzyme

can then be monitored by growth restoration of *E. coli* on a gradient plate that leads to a gradual increase of RNase T expression and activity across the plate. Under these conditions, cells expressing *Eco*CCA can grow up to a distance of 48.5 ± 1.4% of the RNase T gradient, while *Anc*CCA1 expression leads only to cellular growth of up to 37.1 ± 2.5% of the gradient distance, representing a significant difference ($n = 3$; $p = 0.002$). A representative example of the growth behavior of *Eco*CCA and *Anc*CCA1 in this analysis is shown in figure 2*B*. Accordingly, and in agreement with the in vitro turnover numbers, the ancestral candidate is considerably less efficient in catalysis than the extant enzyme.

Substrate Affinity of *Anc*CCA1 Leads to an Increased Reverse Reaction

To characterize the ancestral candidate enzyme in more detail, the binding affinity toward the tRNA substrate was investigated based on an electrophoretic mobility shift assay. Radioactively labeled yeast $tRNA^{Phe}$ ending with CC (substrate) or CCA (product) was incubated with increasing amounts of recombinant *Eco*CCA and *Anc*CCA1, respectively, and separated on a nondenaturing polyacrylamide gel (fig. 3*A*). As expected, the modern enzyme did not show any binding to the substrate, so that no complex formation could be observed over the whole concentration range (0–1.5 µM). According to this weak tRNA affinity, no binding constant could be determined for *Eco*CCA, as already described for several class II CCA-adding enzymes (Shi, Maizels, et al. 1998; Erber et al. 2020; Hennig et al. 2020). The ancestral candidate, however, efficiently bound to the tRNA and showed a strong complex formation. Interestingly, the enzyme did not discriminate between a tRNA that carried an incomplete (tRNA-CC) or a complete CCA end (tRNA-CCA) and formed stable interactions with both transcripts. The robustness of this interaction allowed for the determination of binding constants for both tRNA-CC and tRNA-CCA. For both tRNA versions, a K_D of 300 nM was calculated, comparable to that of many class 1 CCA-adding enzymes in Archaea (Cho et al. 2005, 2008; Shi, Maizels, et al. 1998).

As *Anc*CCA1 efficiently binds to tRNA-CCA as the reaction product of CCA-addition, it is conceivable that this tight interaction promotes a reversal of the polymerization process, leading to tRNA molecules where the CCA end is removed by the enzyme. For class 2 enzymes like *Eco*CCA and *Hsa*CCA (corresponding enzyme from *Homo sapiens*), such a reaction was described, catalyzed in the presence of inorganic pyrophosphate (Igarashi et al. 2011). As this

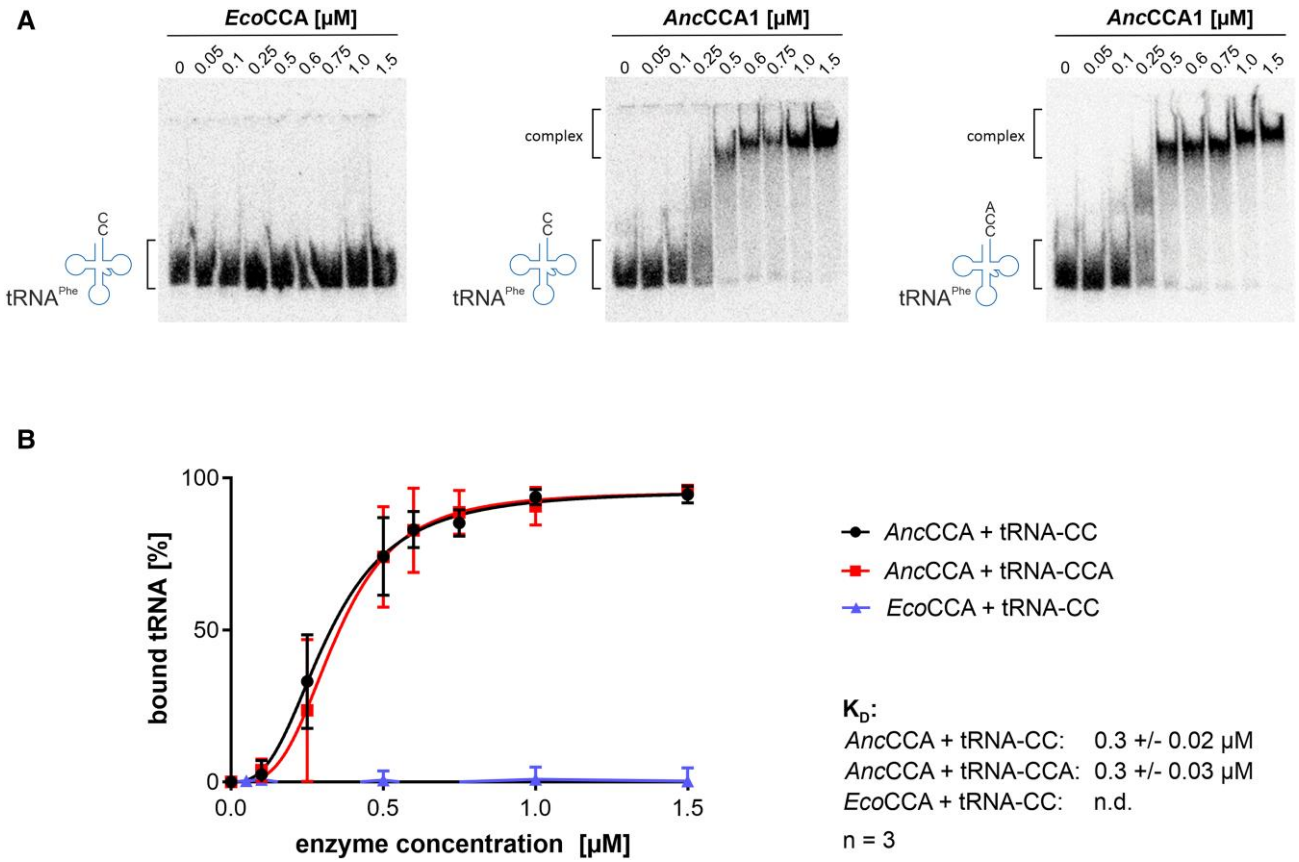

Fig. 3. Binding of *Eco*CCA and *Anc*CCA1 to tRNA. (*A*) In the gel shift experiment, the recombinant *E. coli* enzyme shows no binding to the substrate tRNA-CC, so that is impossible to determine a binding constant. In contrast, *Anc*CCA1 efficiently interacts with a tRNA substrate (tRNA-CC) as well as the final reaction product (tRNA-CCA). (*B*) Based on this strong tRNA binding, dissociation constants of 300 nM for both tRNA-CC and tRNA-CCA could be determined. Data are means ± SD; $n = 3$.

pyrophosphorolysis might interfere with the forward polymerization reaction, it could represent the molecular cause of the observed less efficient CCA-addition catalyzed by the ancestral candidate. To address this question, the pyrophosphorolysis reaction of both *Eco*CCA and *Anc*CCA1 was investigated. Radioactively labeled $tRNA^{Phe}$ ending with CCA was incubated with the recombinant enzymes in the presence of increasing amounts of pyrophosphate (PPi) for 5 min, and the reaction products were separated on a denaturing polyacrylamide gel and visualized by autoradiography (fig. 4A). For *Eco*CCA, a degradation of the CCA end was visible at 0.1 mM PPi that further increased at higher pyrophosphate concentrations. The ancestral enzyme candidate, however, showed a much more pronounced degradation of the tRNA-CCA terminus, starting already at 0.01 mM PPi. In a more detailed analysis, pyrophosphorolysis of both enzymes was investigated depending on enzyme concentration and time (fig. 4B). To obtain comparable results from both enzymes, *Eco*CCA and *Anc*CCA1 activities were normalized based on the forward reaction (CCA-addition). For pyrophosphorolysis, identical units determined for CCA-addition were used. This forward normalization allows for a direct comparison of the efficiency of the reverse reaction. As shown in figure 4B, the ancestral enzyme is much more efficient in CCA degradation than *Eco*CCA. The reverse reaction is immediately visible after 5 min of incubation with 1 forward unit, while for *Eco*CCA, such an effect is only detectable at higher enzyme unit numbers and incubation times. The densitometric analysis exhibits a strong decline of the radioactive signal of tRNA-CCA over time, clearly demonstrating the effectiveness of *Anc*CCA1 in this PPi-dependent CCA end degradation, while *Eco*CCA is much less efficient in its reverse reaction.

Distributive Versus Processive Polymerization

A second possible consequence of the tight substrate binding of *Anc*CCA1 is that this enzyme remains bound to its substrate over the whole CCA-addition reaction, resulting in a processive polymerization mode.

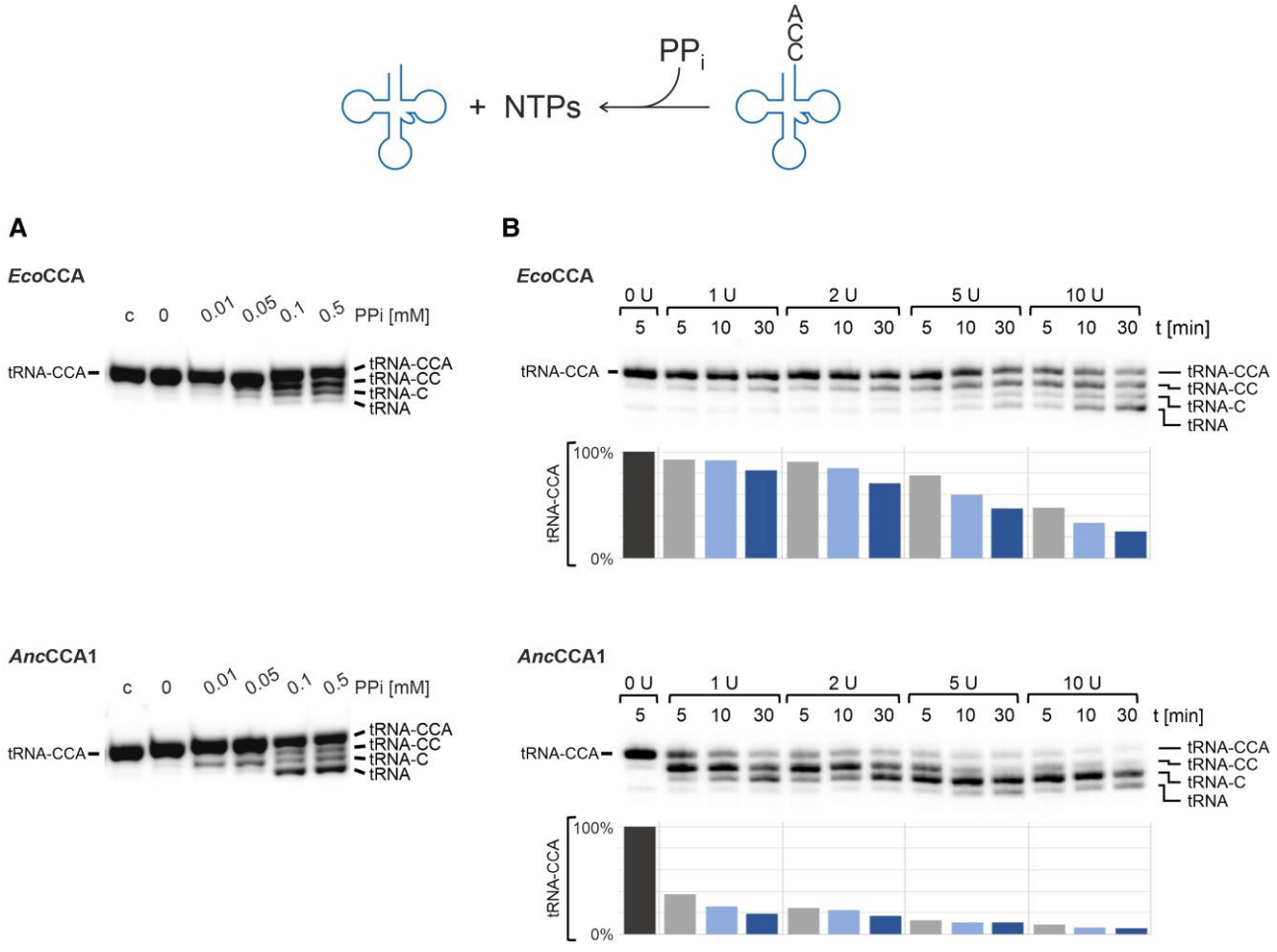

FIG. 4. *Anc*CCA1 efficiently catalyzes the reverse reaction (pyrophosphorolysis). (*A*) In a pyrophosphate concentration series, *Anc*CCA1 degrades the CCA end of a radioactively labeled tRNA already at 0.01 mM PPi, while *Eco*CCA showed only a slight degradation starting at 0.1 mM PPi. (*B*) Pyrophosphorolysis at 1 mM PPi with increasing enzyme concentrations. Both enzyme activities were normalized according to their forward (polymerization) reaction, and equal units were used for pyrophosphorolysis. In the denaturing gel, the decline of the tRNA-CCA band correlates with the increase of lower band intensities, corresponding to tRNAs with partial or no CCA ends. A densitometric analysis of the tRNA-CCA bands clearly shows that *Anc*CCA1 removes CCA ends from the tRNA at high efficiency, even at low enzyme concentrations. In contrast, *Eco*CCA catalyzes the reverse reaction only at higher units and longer incubation times.

Correspondingly, the very low affinity of *Eco*CCA for the tRNA substrate might result in a more distributive reaction, where the enzyme frequently releases the tRNA, even though the CCA end is not yet completed. To identify the polymerization modes of *Eco*CCA and *Anc*CCA1, a pulse-chase experiment was performed (fig. 5). 5 pmol of labeled $tRNA^{Phe}$ substrate were incubated with the corresponding recombinant enzyme for 1, 2, and 5 min in the presence of NTPs. Then, an increasing amount of unlabeled $tRNA^{Phe}$ was added as a competitor to the 1 min incubation assay, and the polymerization reaction was extended for additional 1 or 4 min (in total 2 and 5 min), respectively. After electrophoresis, the *Eco*CCA-catalyzed reaction shows that the addition of the competitor immediately stopped further nucleotide incorporations to the labeled substrate, as the enzyme now binds and adds nucleotides to the excess of competitor (fig. 5A, left panel). Hence, all reaction intermediates (tRNA, tRNA-C, and tRNA-CC) of the labeled substrate remain and are visible as individual bands. During incubation for further 4 min, *Eco*CCA had sufficient time to complete the CCA-addition on labeled as well as unlabeled tRNA, and only higher competitor concentrations (25 and 30 pmol) still distract the enzyme from the labeled substrate, so that reaction intermediates are again visible. The appearance of these reaction intermediates is a clear indication for a distributive polymerization, where the enzyme rapidly falls off its substrate after adding one nucleotide and binds again for further polymerization steps, until CCA-addition is completed (fig. 5A, right panel).

*Anc*CCA1, in contrast, shows a very different behavior (fig. 5*B*, left panel). After 1 min incubation with labeled $tRNA^{Phe}$, only bands corresponding to the tRNA substrate (tRNA without CCA end) and the complete reaction product (tRNA with CCA end) are visible, while reaction intermediates do not appear. The addition of unlabeled competitor tRNA does not change this result, indicating that *Anc*CCA1 tightly binds to its substrate and adds a complete CCA end before dissociating from the product tRNA-CCA. Such a behavior represents a typical processive polymerization reaction. Hence, while both enzymes show a similar specificity and fidelity of CCA-addition, they differ dramatically in their substrate affinity, and, as a consequence, in the polymerization mode and the corresponding efficiency of the reverse reaction (fig. 6).

Discussion

As an essential feature of functional tRNA molecules, the CCA end represents the site of aminoacylation. In most organisms, this triplet is not encoded in the corresponding tRNA genes and has to be added posttranscriptionally by a highly specialized RNA polymerase, the CCA-adding enzyme (tRNA nucleotidyltransferase). *E. coli* is an exception

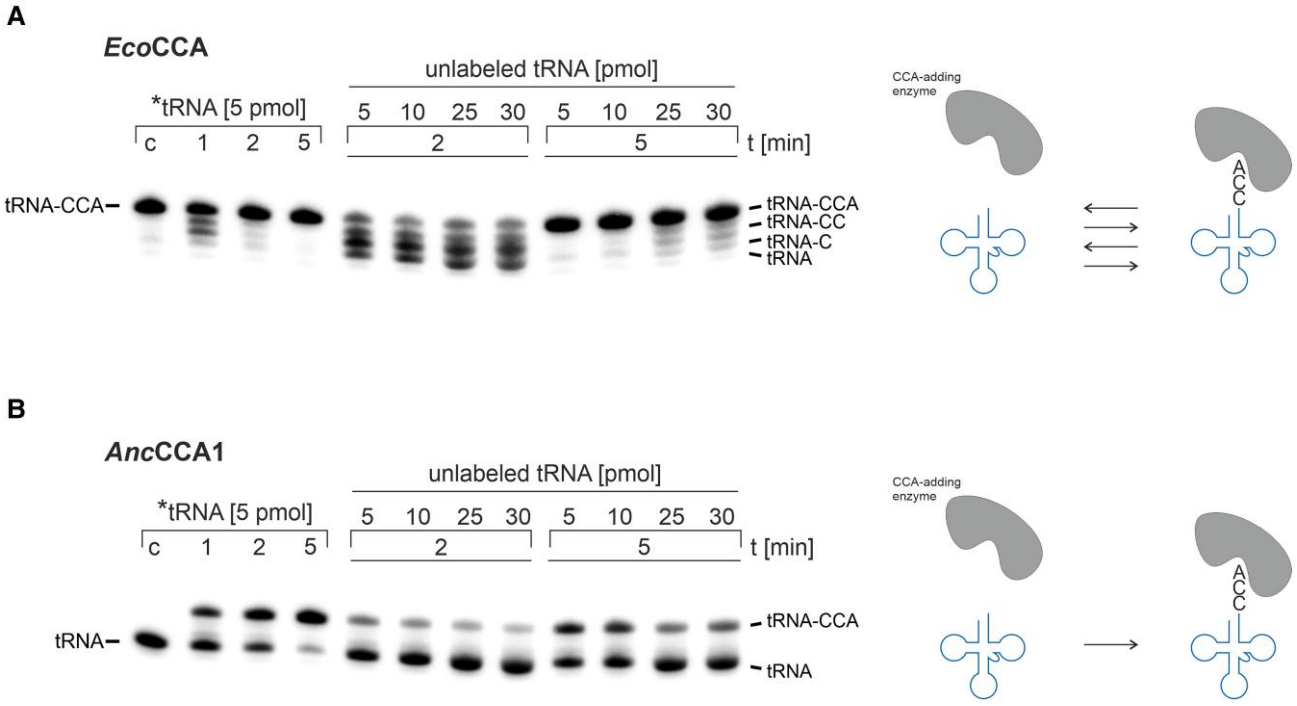

Fig. 5. *Eco*CCA and *Anc*CCA1 differ in their polymerization modes. (*A*) During 1 min incubation with labeled tRNA (*tRNA) and NTPs, *Eco*CCA adds individual nucleotides to the substrate, resulting in a band pattern that corresponds to the individual reaction intermediates. After addition of increasing amounts of unlabeled competitor tRNA, the enzyme is distracted from the labeled substrate, so that these intermediates are not completed and remain visible, a typical indication of distributive polymerization. (*B*) *Anc*CCA1 shows no reaction intermediates during polymerization. The presence of the competitor tRNA in excess does not lead to the release of tRNA molecules with incomplete CCA ends. Obviously, the enzyme remains bound to the tRNA substrate until CCA synthesis is completed, corresponding to a processive polymerization reaction.

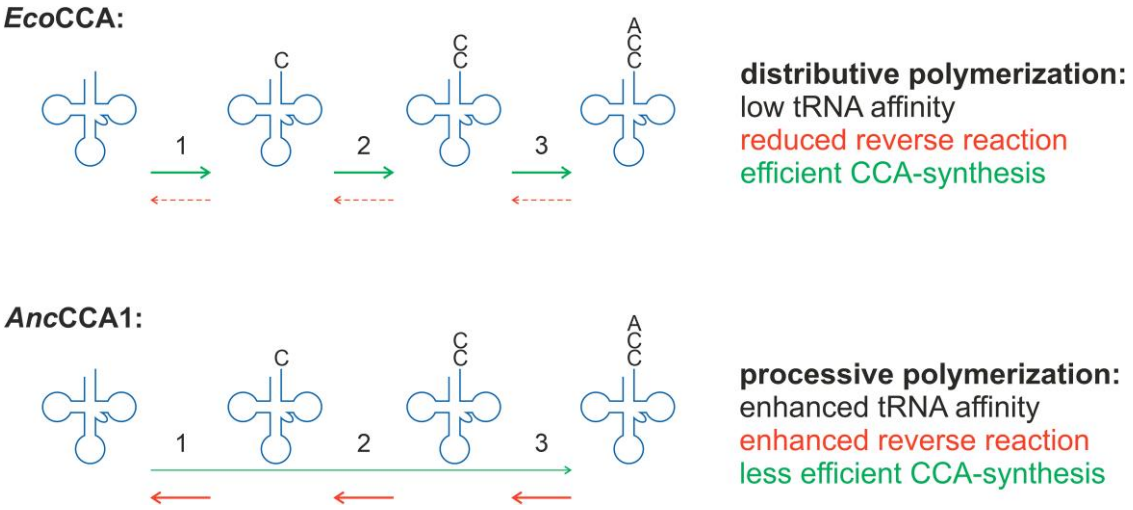

Fig. 6. Consequences of distributive versus processive polymerization as a result of differences in substrate affinity. Despite a low substrate affinity, the extant *E. coli* enzyme shows an efficient CCA synthesis based on a distributive polymerization mode. In contrast, the ancestral candidate *Anc*CCA1 acts as a processive polymerase, as it tightly binds its tRNA substrate. As a consequence, it efficiently promotes the reverse reaction, leading to an overall less efficient CCA synthesis.

from this rule, as all tRNA genes encode for the CCA end. Yet, this organism expresses a typical CCA-adding enzyme, and while its deletion is not lethal, it leads to a substantial growth restraint, as the enzyme is also required for repair and proper maintenance of the CCA ends (Zhu and Deutscher 1987). In addition, it is involved in tRNA quality control (Wellner et al. 2018; Wilusz et al. 2011). While nucleotide selectivity as well as reaction mechanism are understood in quite some detail (Xiong and Steitz 2006; Betat et al. 2010; Tomita and Yamashita 2014), the bacterial and eukaryotic enzymes exhibit a highly puzzling and at a first glance contradicting feature. For an efficient polymerization reaction, one would expect that these enzymes show a considerable affinity to their tRNA substrates. Yet, this is not the case—the substrate interaction is so weak that no binding constants could be determined so far. Only enzymes that are specialized to aberrant tRNA structures or that carry Q/N repeats of unknown function that might contribute to tRNA binding show a detectable affinity to tRNAs (Erber et al. 2020; Hennig et al. 2020).

To clarify how and why modern CCA-adding enzymes evolved such conflicting properties, ASR was utilized to obtain approximations of ancestral CCA-adding enzymes that can be compared with their extant counterparts. With a focus on protein sequences from Gammaproteobacteria, the resulting phylogenetic tree led to the reconstruction of four individual ancestral enzyme candidates with efficient CCA-adding activity, where *Anc*CCA1 was selected for an in-depth analysis (supplementary fig. S3, Supplementary Material online). A first deviation from the extant *Eco*CCA is that this reconstruction exhibits a less efficient catalysis of CCA-addition, although it shows a polymerization fidelity indistinguishable from that of the extant enzyme. A reason for this reduced efficiency is the high tRNA substrate affinity of *Anc*CCA1, as this tight binding results in an increased reverse reaction, where the CCA end is partially or completely removed from the tRNA. As a consequence, CCA synthesis runs into a futile cycle of polymerization and degradation (fig. 6). In addition, the enzyme does not discriminate between substrate (tRNA without CCA end) and product (tRNA with CCA end), as it has identical affinities to both types of tRNA. Hence, a second parameter reducing the polymerization efficiency is an inefficient product release, as an enzyme bound to tRNA-CCA is blocked for CCA-addition on another tRNA molecule, and our results demonstrate that lower enzymatic efficiency is visible both in vitro as well as in vivo.

In *Eco*CCA, however, efficient substrate binding was obviously lost in evolution in order to avoid these detrimental side effects of reverse reaction and product inhibition. As a consequence, the extant enzyme frequently dissociates from its tRNA substrate and represents a distributive, not a processive polymerase. Yet, the catalytic activity of the extant enzyme is superior to that of the ancestral candidate.

Interestingly, several archaeal-type CCA-adding enzymes (comprising class 1) show similar or even better binding affinities to the tRNA substrate while still catalyzing the CCA-addition at high efficiency (Shi, Maizels, et al. 1998; Okabe et al. 2003; Cho et al. 2005, 2008). This obvious discrepancy to the situation of class 2 enzymes can be explained by the fact that class 1 enzymes do not catalyze a reverse reaction that would degrade the CCA terminus and, consequently, reduce the catalytic efficiency (Igarashi et al. 2011). How these enzymes manage to avoid the reverse reaction despite a high affinity toward the tRNA is not known. A crystal structure of the *Archaeoglobus fulgidus* enzyme bound to a tRNA-mimicking RNA minihelix ending with CCA shows that the overall structure is not different to the situation where reaction intermediates (minihelix with partial CCA ends) are bound (Tomita et al. 2006). Yet, it

appears that this structural state corresponds to the situation immediately after CCA-addition, when the pyrophosphate still is present in the catalytic core. A second co-crystal structure of the same enzyme reveals a more open conformation that seems to represent the state after release of the pyrophosphate (Xiong and Steitz 2004). It is possible that these two structures indicate the conformational dynamics that allow the enzyme to get rid of the pyrophosphate in order to promote the forward reaction (CCA synthesis). Why class 2 CCA-adding enzymes follow a different strategy of reduced substrate binding and not the more obvious and straightforward efficient release of one of the reaction products remains a mystery. Here, detailed high-resolution crystal or cryo-EM structures of the corresponding reaction states are required. Nevertheless, it is fascinating to see that enzymes catalyzing identical and essential reactions find different solutions to solve the same problem, although it comes as a big surprise that a significant lowering of substrate binding leading to a distributive polymerization mode results in a considerable increase in polymerization efficiency. Hence, this comparative analysis of enzymatic features demonstrates an impressive potential of ASR as a tool to understand an otherwise confusing catalytic behavior of a modern RNA polymerase.

Materials and Methods

Sequences

Sequences for CCA-adding enzymes were taken from the NCBI database (ncbi.nlm.nih.gov/assembly/), which included 43,865 gammaproteobacterial RefSeq entries and 3,387 complete gammaproteobacterial genomes in December 2017. Genome assemblies and assembly structure reports of 272 representatives of Gammaproteobacteria were downloaded. Using Blastp (blast.ncbi.nlm.nih.gov/Blast.cgi? PAGE = Proteins) and a small set of known CCA-adding enzyme sequences from Gammaproteobacteria, sequences of CCA-adding enzymes were identified in the corresponding proteomes.

Alignment and Phylogenetic Tree Reconstruction

Alignment was carried out using Clustal Omega (Sievers and Higgins 2018). To verify the reliability of the alignment, the tool CORE (Notredame et al. 2000) was applied to the pre-computed alignment. The CORE (Consistency of the Overall Residue Evaluation) index indicates the consistency for each residue on an integer scale from 9 to 0, respectively. 75% of residues had the highest CORE index, while only 3% were considered to be not properly aligned (CORE indices 0, 1, and 3). In addition to the higher number of gaps in the 17 shortest sequences, 10% of their residues were categorized as improperly aligned.

For phylogenetic reconstruction, columns with more than 50% gaps were removed with GBLOCKS v.0.91b, resulting in the retention of 347 of 470 alignment positions (73%) (Castresana 2000; Talavera and Castresana 2007). The gene tree was computed using the maximum likelihood approach implemented in RAxML (Stamatakis 2014) and the following parameter settings: LG substitution model with gamma-distribution site rates and 100 bootstrap replicates (Le and Gascuel 2008).

Reconstruction of Ancestral Sequences

For computation of the ancestral state at the root of the gene tree, FastML was used (Pupko et al. 2000). This algorithm uses a maximum likelihood method and offers character reconstruction based on the computation of joint and marginal probabilities as well as indel reconstruction, a way to model the evolutionary insertion-deletion events. The joint and marginal reconstructions gave highly similar results, showing 3% of substitutions. With an indel cutoff of 0.9, that is only highly likely indels are considered, the sequences contained six indels, four smaller ones (1–2 aa) and two of 4 and 6 aa. At a cutoff of 0.7 or lower, a seventh indel of 5 aa became apparent (supplementary fig. S2, Supplementary Material online). Comparisons revealed that all indel positions were associated with very low values for alignment consistency. Such a correlation between measures of alignment quality and reconstruction accuracy has been previously described (Vialle et al. 2018). For functional testing, the following reconstructed sequences were used: the shortest sequence (*Anc*CCA1) resulting from joint reconstruction and indel cutoff 0.7, the joint reconstruction without indels (*Anc*CCA2) and the joint and marginal reconstructions with indel cutoff 0.9 (*Anc*CCA3 and *Anc*CCA4, respectively) (supplementary fig. S3, Supplementary Material online).

To validate the robustness of the sequence reconstruction with respect to taxon sampling, we composed seven new sets by removing or adding specific sequences to the original set and computed ancestral sequences the same way as for *Anc*CCA1 (supplementary table S2, Supplementary Material online). After alignment of the ancestral sequences (*Anc*CCA1 and ASR1–7) to the full multiple sequence alignment, a distance matrix was computed and visualized as a heatmap (supplementary fig. S5, Supplementary Material online). The obtained distance matrix is in good agreement with the phylogenetic gene tree. The reconstructed ancestral sequences are very similar to each other (with a highest value of 20.0) and more distant to all extant sequences (with a minimum distance of 28.2). Due to the pectinate tree topology, Xanthomonadales split from all other orders at the ancestral node. Distances between Xanthomonadales and any other order should therefore be larger than their distances to the reconstructed ancestors. This was verified by looking at triples composed of the set of reconstructed sequences and sequences from Xanthomonadales (Xan) and either Enterobacteriales (Ent) or Chromatiales (Chr). In both cases, the mean distances of the ancestral reconstructions to sequences of the extant orders were smaller than the distances between the extant orders themselves (ASRs—Xan: 44.09, ASRs—Ent: 54.18, Xan—Ent: 75.47,

ASRs—Chr: 40.15, Xan—Chr: 64.86; supplementary fig. S6, Supplementary Material online).

Cloning, Expression and Purification of Recombinant CCA-Adding Enzymes

Codon-optimized open reading frames encoding the candidates for the ancestral gammaproteobacterial CCA-adding enzyme were synthesized by GeneScript (Rijswijk, Netherlands) and inserted into pET28a(+) using *Xba*I/*Xho*I restriction sites. The ORFs were fused to a sequence encoding a C-terminal His_6 tag. The ORF for the *E. coli* CCA-adding enzyme was cloned in an identical way into the same vector system. Recombinant proteins were expressed in *E. coli* BL21 (DE3) cca::cam lacking the endogenous CCA-adding enzyme and purified as described (Ernst et al. 2018). Purity of individual fractions was monitored by SDS-PAGE electrophoresis followed by staining with Coomassie Brilliant Blue R-250 (BioRad). Enzyme preparations were stored in the presence of 10% (w/v) glycerol at −80°C until use.

Preparation of RNA Substrates

In vitro transcribed tRNA substrates were prepared as described (Schürer et al. 2002; Mörl and Hartmann 2014). For internal labeling, α-^{32}P-ATP (3,000 Ci/mmol) was added to the transcription reaction.

In Vitro CCA-Incorporation and Determination of Arbitrary Units

Nucleotide incorporation assays were performed as described (Wende et al. 2015; Ernst et al. 2018). Sequence analysis of reaction products was carried out according to Wende et al. (2015). For the definition of arbitrary units, enzyme activity was determined on the standard substrate $tRNA^{Phe}$ from *Saccharomyces cerevisiae* as described (Hennig et al. 2020). An arbitrary unit was defined as the enzyme amount leading to 50% substrate turnover in 30 min.

Kinetic Analysis

Michaelis–Menten kinetic parameters for CC- and A-addition were determined for *Eco*CCA and *Anc*CCA1 as described (Just et al. 2008; Hennig et al. 2020) in three independent experiments and calculated by nonlinear regression (GraphPadPrism 7). Due to limited solubility, tRNA transcripts cannot be used at excessive saturating conditions. Hence, the calculated kinetic parameters represent apparent values, as frequently used for this type of enzymes (Tomita et al. 2004, 2006; Hoffmeier et al. 2010; Wende et al. 2015).

In Vivo Complementation

Determination of in vivo activity of *Eco*CCA and *Anc*CCA1 was carried out according to Wellner et al. (2019).

Electrophoretic Mobility Shift Assay

tRNA binding constants of *Eco*CCA and *Anc*CCA1 were determined according to Hennig et al. (2020). For each enzyme, dissociation constants of three independent experiments were calculated using GraphPadPrism 7.

Pyrophosphorolysis

Analysis of CCA end degradation in the presence of increasing amounts of KPPi was performed according to Igarashi et al. (2011). Time and enzyme concentration series of pyrophosphorolysis were performed at 37°C in standard CCA-addition buffer (30 mM HEPES/KOH pH 7.6, 30 mM KCl, 6 mM $MgCl_2$, 2 mM DTT), 5 pmol ^{32}P-labeled $tRNA^{Phe}$ + CCA and 0–10 units CCA-adding enzyme (enzyme activities were normalized as arbitrary units according to the forward reaction). KPPi was added to a final concentration of 1 mM. Reaction was stopped by ethanol precipitation, and reaction products were separated on denaturing 10% polyacrylamide gels and visualized by autoradiography. Densitometric analysis was done with ImageQuant TL.

Pulse-Chase Analysis of CCA-Addition

Pulse-chase reactions were incubated at 37°C and contained 30 mM HEPES/KOH pH 7.6, 30 mM KCl, 6 mM $MgCl_2$, 2 mM DTT, 1 mM NTPs, and 5 ng of *Eco*CCA or 30 ng of *Anc*CCA1, respectively. Pulse reaction was initiated by the addition of ^{32}P-labeled $tRNA^{Phe}$ for 1 min. The chase was started by adding 5–30 pmol of unlabeled $tRNA^{Phe}$ as competitor and further incubation for 1 and 4 min. Reaction products were ethanol precipitated, separated on denaturing 10% polyacrylamide gels and visualized by autoradiography.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) [MO 634/8-2 and PR 1288/6-2]. We acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) and Leipzig University within the program of Open Access Publishing. We thank Rainer Merkl, Reinhard Sterner, and Kristina Straub for helpful discussion concerning ancestral sequence reconstruction, Alexandra Bluhm and Gina Münch for experimental support. Thanks to Sonja Bonin and Tobias Friedrich for expert technical assistance.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Data Availability

Selected gammaproteobacterial orders and data base entries of coding sequences for ancestral sequence reconstruction of CCA-adding enzymes are available in the Supplementary material.

References

Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T. 2012. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res.* **40**: W580–W584.

Battistuzzi FU, Feijao A, Hedges SB. 2004. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol.* **4**:44.

Betat H, Rammelt C, Martin G, Mörl M. 2004. Exchange of regions between bacterial poly(A) polymerase and the CCA-adding enzyme generates altered specificities. *Mol Cell.* **15**:389–398.

Betat H, Rammelt C, Mörl M. 2010. tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cell Mol Life Sci.* **67**:1447–1463.

Carrigan MA, Uryasev O, Frye CB, Eckman BL, Myers CR, Hurley TD, Benner SA. 2015. Hominids adapted to metabolize ethanol long before human-directed fermentation. *Proc Natl Acad Sci USA* **112**:458–463.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* **17**: 540–552.

Cho HD, Sood VD, Baker D, Weiner AM. 2008. On the role of a conserved, potentially helix-breaking residue in the tRNA-binding α-helix of archaeal CCA-adding enzymes. *RNA* **14**:1284–1289.

Cho HD, Verlinde CL, Weiner AM. 2005. Archaeal CCA-adding enzymes: central role of a highly conserved beta-turn motif in RNA polymerization without translocation. *J Biol Chem.* **280**: 9555–9566.

Cudny H, Deutscher MP. 1986. High-level overexpression, rapid purification, and properties of Escherichia coli tRNA nucleotidyltransferase. *J Biol Chem.* **261**:6450–6453.

Cudny H, Lupski JR, Godson GN, Deutscher MP. 1986. Cloning, sequencing, and species relatedness of the *Escherichia coli cca* gene encoding the enzyme tRNA nucleotidyltransferase. *J Biol Chem.* **261**:6444–6449.

Erber L, Hoffmann A, Fallmann J, Hagedorn M, Hammann C, Stadler PF, Betat H, Prohaska S, Mörl M. 2020. Unusual occurrence of two bona-fide CCA-adding enzymes in *Dictyostelium discoideum*. *Int J Mol Sci.* **21**:15.

Ernst FGM, Erber L, Sammler J, Jühling F, Betat H, Mörl M. 2018. Cold adaptation of tRNA nucleotidyltransferases: a tradeoff in activity, stability and fidelity. *RNA Biol.* **15**:144–155.

Ernst FGM, Rickert C, Bluschke A, Betat H, Steinhoff H-J, Mörl M. 2015. Domain movements during CCA-addition: a new function for motif C in the catalytic core of the human tRNA nucleotidyltransferases. *RNA Biol.* **12**:435–446.

Hennig O, Philipp S, Bonin S, Rollet K, Kolberg T, Jühling T, Betat H, Sauter C, Mörl M. 2020. Adaptation of the *Romanomermis culicivorax* CCA-adding enzyme to miniaturized armless tRNA substrates. *Int J Mol Sci.* **21**:23.

Hoffmeier A, Betat H, Bluschke A, Günther R, Junghanns S, Hofmann H-J, Mörl M. 2010. Unusual evolution of a catalytic core element in CCA-adding enzymes. *Nucleic Acids Res.* **38**:4436–4447.

Hou YM. 2000. Unusual synthesis by the *Escherichia coli* CCA-adding enzyme. *RNA.* **6**:1031–1043.

Igarashi T, Liu C, Morinaga H, Kim S, Hou Y-M. 2011. Pyrophosphorolysis of CCA addition: implication for fidelity. *J Mol Biol.* **414**:28–43.

Jermann TM, Opitz JG, Stackhouse J, Benner SA. 1995. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* **374**:57–59.

Just A, Butter F, Trenkmann M, Heitkam T, Mörl M, Betat H. 2008. A comparative analysis of two conserved motifs in bacterial poly(A) polymerase and CCA-adding enzyme. *Nucleic Acids Res.* **36**:5212–5220.

Kim S, Liu C, Halkidis K, Gamper HB, Hou Y-M. 2009. Distinct kinetic determinants for the stepwise CCA addition to tRNA. *RNA* **15**: 1827–1836.

Kumar S, Stecher G, Suleski M, Hedges SB. 2017. Timetree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol.* **34**: 1812–1819.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol.* **25**:1307–1320.

Li F, Xiong Y, Wang J, Cho HD, Tomita K, Weiner AM, Steitz TA. 2002. Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell* **111**:815–824.

Maizels N, Weiner AM. 1994. Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc Natl Acad Sci USA* **91**:6729–6734.

Maizels N, Weiner AM. 1999. The genomic tag hypothesis: what molecular fossils tell us about the evolution of tRNA synthesis. In: *The RNA world*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. p. 79–111.

Maizels N, Weiner AM, Yue D, Shi PY. 1999. New evidence for the genomic tag hypothesis: archaeal CCA-adding enzymes and tDNA substrates. *Biol Bull.* **196**:331–333; discussion 333–4.

Merkl R, Sterner R. 2016a. Ancestral protein reconstruction. Techniques and applications. *Biol Chem.* **397**:1–21.

Merkl R, Sterner R. 2016b. Reconstruction of ancestral enzymes. *Perspect Sci.* **9**:17–23.

Mörl M, Hartmann RK. 2014. Production of RNAs with homogeneous 5′- and 3′-ends. In: Hartmann RK Bindereif A Schön A and Westhof E, editors. *Handbook of RNA biochemistry*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA. p. 29–44.

Neuenfeldt A, Just A, Betat H, Mörl M. 2008. Evolution of tRNA nucleotidyltransferases: a small deletion generated CC-adding enzymes. *Proc Natl Acad Sci USA* **105**:7953–7958.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol.* **302**:205–217.

Okabe M, Tomita K, Ishitani R, Ishii R, Takeuchi N, Arisaka F, Nureki O, Yokoyama S. 2003. Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J.* **22**:5918–5927.

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, *et al.* 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**:D733–D745.

Pauling L, Zuckerkandl E, Henriksen T, Lövstad R. 1963. Chemical paleogenetics. Molecular "restoration studies" of extinct forms of life. *Acta Chem Scand.* **17**:9–16.

Peltier MR, Raley LC, Liberles DA, Benner SA, Hansen PJ. 2000. Evolutionary history of the uterine serpins. *J Exp Zool.* **288**: 165–174.

Pupko T, Pe'er I, Shamir R, Graur D. 2000. A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol Biol Evol.* **17**:890–896.

Risso VA, Manssour-Triedo F, Delgado-Delgado A, *et al.* 2015. Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol Biol Evol.* **32**:440–455.

Sánchez-Ruiz J, Risso VA. 2018. Ancestral proteins: how and why. *Biofísica* **12**:13–18.

Schimmel P, Yang X-L. 2004. Two classes give lessons about CCA. *Nat Struct Mol Biol.* **11**:807–808.

Schürer H, Lang K, Schuster J, Mörl M. 2002. A universal method to produce in vitro transcripts with homogeneous 3′ ends. *Nucleic Acids Res.* **30**:e56.

Shi PY, Maizels N, Weiner AM. 1998. CCA Addition by tRNA nucleotidyltransferase: polymerization without translocation? *EMBO J.* **17**:3197–3206.

Shi PY, Weiner AM, Maizels N. 1998. A top-half tDNA minihelix is a good substrate for the eubacterial CCA-adding enzyme. *RNA* **4**: 276–284.

Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**:135–145.

Stackhouse J, Presnell SR, McGeehan GM, Nambiar KP, Benner SA. 1990. The ribonuclease from an extinct bovid ruminant. *FEBS Lett.* **262**:104–106.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol.* **56**:564–577.

Thomson JM, Gaucher EA, Burgan MF, De Kee DW, Li T, Aris JP, Benner SA. 2005. Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat Genet.* **37**:630–635.

Toh Y, Takeshita D, Numata T, Fukai S, Nureki O, Tomita K. 2009. Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. *EMBO J.* **28**:3353–3365.

Tomita K, Fukai S, Ishitani R, Ueda T, Takeuchi N, Vassylyev DG, Nureki O. 2004. Structural basis for template-independent RNA polymerization. *Nature* **430**:700–704.

Tomita K, Ishitani R, Fukai S, Nureki O. 2006. Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature* **443**:956–960.

Tomita K, Yamashita S. 2014. Molecular mechanisms of template-independent RNA polymerization by tRNA nucleotidyltransferases. *Front Genet.* **5**:Article 36.

Tretbar S, Neuenfeldt A, Betat H, Mörl M. 2011. An inhibitory C-terminal region dictates the specificity of A-adding enzymes. *Proc Natl Acad Sci USA* **108**:21040–21045.

Vialle RA, Tamuri AU, Goldman N. 2018. Alignment modulates ancestral sequence reconstruction accuracy. *Mol Biol Evol.* **35**:1783–1797.

Vörtler S, Mörl M. 2010. tRNA-nucleotidyltransferases: highly unusual RNA polymerases with vital functions. *FEBS Lett.* **584**:297–302.

Wallace IM, O'Sullivan O, Higgins DG, Notredame C. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* **34**:1692–1699.

Weiner AM, Maizels N. 1999. The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication, and clues regarding the origin of protein synthesis. *Biol Bull.* **196**:327–330.

Wellner K, Czech A, Ignatova Z, Betat H, Mörl M. 2018. Examining tRNA 3′-ends in *Escherichia coli*: teamwork between CCA-adding enzyme, RNase T, and RNase R. *RNA* **24**:361–370.

Wellner K, Pöhler M-T, Betat H, Mörl M. 2019. Dual expression of CCA-adding enzyme and RNase T in *Escherichia coli* generates a distinct CCA growth phenotype with diverse applications. *Nucleic Acids Res.* **47**:3631–3639.

Wende S, Bonin S, Götze O, Betat H, Mörl M. 2015. The identity of the discriminator base has an impact on CCA addition. *Nucleic Acids Res.* **43**:5617–5629.

Williams KP, Gillespie JJ, Sobral BWS, Nordberg EK, Snyder EE, Shallom JM, Dickerman AW. 2010. Phylogeny of Gammaproteobacteria. *J Bacteriol.* **192**:2305–2314.

Wilusz JE, Whipple JM, Phizicky EM, Sharp PA. 2011. tRNAs marked with CCACCA are targeted for degradation. *Science* **334**: 817–821.

Xiong Y, Steitz TA. 2004. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430**:640–645.

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.

Zhu L, Deutscher MP. 1987. tRNA nucleotidyltransferase is not essential for Escherichia coli viability. *EMBO J.* **6**:2473–2477.