Cellular and Molecular Life Sciences

REVIEW

tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization

Heike Betat · Christiane Rammelt ·
Mario Mörl

Abstract RNA polymerases are important enzymes involved in the realization of the genetic information encoded in the genome. Thereby, DNA sequences are used as templates to synthesize all types of RNA. Besides these classical polymerases, there exists another group of RNA polymerizing enzymes that do not depend on nucleic acid templates. Among those, tRNA nucleotidyltransferases show remarkable and unique features. These enzymes add the nucleotide triplet C–C–A to the 3′-end of tRNAs at an astonishing fidelity and are described as "CCA-adding enzymes". During this incorporation of exactly three nucleotides, the enzymes have to switch from CTP to ATP specificity. How these tasks are fulfilled by rather simple and small enzymes without the help of a nucleic acid template is a fascinating research area. Surprising results of biochemical and structural studies allow scientists to understand at least some of the mechanistic principles of the unique polymerization mode of these highly unusual enzymes.

H. Betat · M. Mörl (✉)
Institute for Biochemistry, University of Leipzig,
Brüderstr. 34, 04103 Leipzig, Germany
e-mail: moerl@uni-leipzig.de

C. Rammelt
Institute for Biochemistry,
Martin Luther University Halle-Wittenberg,
Kurt-Mothes-Str. 3, 06120 Halle, Germany

Introduction

The transfer of nucleotides onto acceptor molecules is a frequent reaction in various biological processes. The enzymes catalyzing this type of reaction belong to a huge class of nucleotidyltransferases that add nucleotides to substrates like nucleic acids, proteins, or antibiotics, and are involved in a variety of reactions such as DNA replication, telomere maintenance, transcription, DNA/RNA ligation, and RNA processing [1–3].

According to sequence and structural homologies, nucleotidyltransferases are subdivided into distinct superfamilies. A prominent member of such a subdivision is the eukaryotic DNA polymerase β, one of the smallest DNA polymerases identified so far. Phylogenetic analyses revealed a series of other nucleotidyltransferases that share a common sequence signature motif and, consequently, the mechanistic basis of nucleotidyl transfer with this DNA polymerase [1, 2]. Summarized as the polymerase β family, these enzymes show a surprisingly broad substrate and reaction spectrum that is not found in other nucleotidyl-transferase superfamilies [4, 5].

The sequence motif common to all members of this polymerase β superfamily consists of the amino acid signature hG[GS]x(9,13)Dh[DE]h (x represents any amino acid, h represents a hydrophobic amino acid) [2]. This motif is part of the consensus fold, consisting of a five-stranded antiparallel β sheet that is flanked by two α helices [2]. The β sheet contains two highly conserved carboxylates (DxD or DxE) that coordinate two essential metal ions involved in catalysis.

The polymerase β signature element is also present in a series of small proteins of less than 100 amino acids that were identified in archaea and bacteria. Interestingly, these minimal nucleotidyltransferases (MNTs) carry no

additional functional domains and probably depend in their activity on the interaction with other regulatory or specificity subunits. Accordingly, it was speculated that modern nucleotidyltransferases evolved by fusing MNT-like core enzymes with new domains for substrate binding, regulation, and further activities [1, 6].

Besides the common signature motif described above, further sequence elements are used to divide the polymerase β superfamily into two subgroups that share only little sequence homology besides the signature mentioned above. Class I consists of polymerase β and other members of the polymerase X family, terminal nucleotidyltransferases (TdT), antibiotic nucleotidyltransferases, protein nucleotidyltransferases, $2'–5'$-oligo(A) synthetases, Trf4-like poly(A) polymerases, eukaryotic poly(A) polymerases, archaeal tRNA nucleotidyltransferases, and many others. Class II, on the other hand, is a rather small group of bacterial and eukaryotic tRNA nucleotidyltransferases and bacterial poly(A) polymerases [1, 5, 6]. Whereas class I enzymes are rather heterogeneous and do not exhibit any further sequence similarity, the members of class II share a highly homologous N-terminal region of about 25 kDa with several conserved active site motifs. Phylogenetic analyses indicate that the two classes are only distantly related and separated rather early in evolution [1, 5].

Poly(A) polymerase, terminal uridylyl transferase, terminal deoxynucleotidyltransferase, and tRNA nucleotidyltransferase represent a specific subgroup of these two classes, as they add nucleotides to a DNA or RNA primer without using a nucleic acid template. Whereas terminal deoxynucleotidyltransferase adds all four types of deoxynucleotides to DNA $3'$-ends and promotes antigen receptor diversity [7], the other three enzymes play key roles in the processing of RNA transcripts, adding specific nucleotides to RNA $3'$-ends. Poly(A) polymerases synthesize long homopolymeric poly-adenylate tails at the $3'$-end of mRNAs (in some cases also on tRNAs and other transcripts) [8], while the RNA editing enzyme uridylyl transferase adds UMP residues to mRNAs and several small transcripts like gRNAs and U2 snRNA [9, 10].

A highly sophisticated enzyme in this group, however, is tRNA nucleotidyltransferase. This protein represents an important tRNA processing enzyme and synthesizes and maintains the specific sequence C–C–A at the $3'$-end of tRNAs ("CCA-adding enzyme"). After additional processing steps, the resulting mature tRNA molecules play a central role in protein biosynthesis, acting as adapter molecules for translating the genetic information stored in the genome into protein sequences [11–14]. During this process, tRNAs are charged with their cognate amino acid and deliver their cargo to the ribosome according to the mRNA sequence to be translated. In these events, the $3'$-terminal CCA end of the tRNA (added by tRNA nucleotidyltransferase) plays an essential role. First, the CCA end is the site of aminoacylation, and aminoacyl tRNA synthetases fuse the individual amino acids to the ribose moiety of the terminal A residue [15]. Second, the CCA terminus is required for the correct positioning of the aminoacyl-tRNA in the ribosome's A- and P-site in order to guarantee an efficient peptidyl transfer reaction [16]. Furthermore, the CCA sequence is involved in translation termination when the nascent peptide is released from the ribosome. Here, the CCA triplet is important for the coordination of water molecules hydrolyzing the ester linkage at the peptidyl-tRNA [17].

Despite these vital functions of the CCA terminus, very few organisms encode this sequence in the tRNA genes, while in most species, this triplet has to be added posttranscriptionally [5, 18]. Consequently, the CCA-adding enzyme represents an essential activity in the majority of organisms. In *E. coli*, on the other hand, where CCA ends are encoded, this enzyme is dispensable, and a corresponding gene knockout is not lethal. Yet, the phenotype is a dramatic growth impairment, indicating the repair function of the CCA-adding enzyme on defective tRNAs lacking CCA ends due to hydrolytic damage [19–21].

Hence, the CCA-adding enzymes represent vital components of the cell's tRNA maturation and maintenance system. Due to their extraordinary mechanism of nucleotide selection and incorporation, these enzymes are some of the most fascinating nucleic acid polymerases.

Upon a closer look, the simple addition of two C and one A residue catalyzed by the CCA-adding enzymes turns out to be a rather complicated event. These enzymes have to solve a series of tricky problems. First, a CCA-adding enzyme has to recognize tRNA and tRNA-like structures as substrates; second, the correct nucleotides CTP and ATP have to be selected and discriminated against UTP and GTP; third, after incorporation of two C residues, the nucleotide specificity has to switch towards ATP without the help of a nucleic acid template; fourth, the enzymes have to stop polymerization exactly after three positions (many other polymerases are rather sloppy in ending nucleotide addition [22, 23]); and fifth, the enzymes do not stubbornly add CCA-ends to their substrates, but recognize partial CCA-ends and add only the missing residues for completion.

Before crystal structures of CCA-adding enzymes became available, biochemical and enzymological studies as well as in vivo experiments led to the first—and in many cases correct—ideas about the underlying reaction mechanisms [24–26]. In the last decade, however, crystal structures as well as novel biochemical approaches led to more detailed insights into architecture and reaction mechanisms of these enzymes as well as their evolution

[27]. Based on these new and surprising findings, we focus in this review on the unique features and evolutionary pathways of CCA-adding enzymes—essential RNA polymerases that emerged twice in evolution—leading to different structural characteristics and unusual mechanistic solutions for an error-free and sequence-specific CCA polymerization reaction.
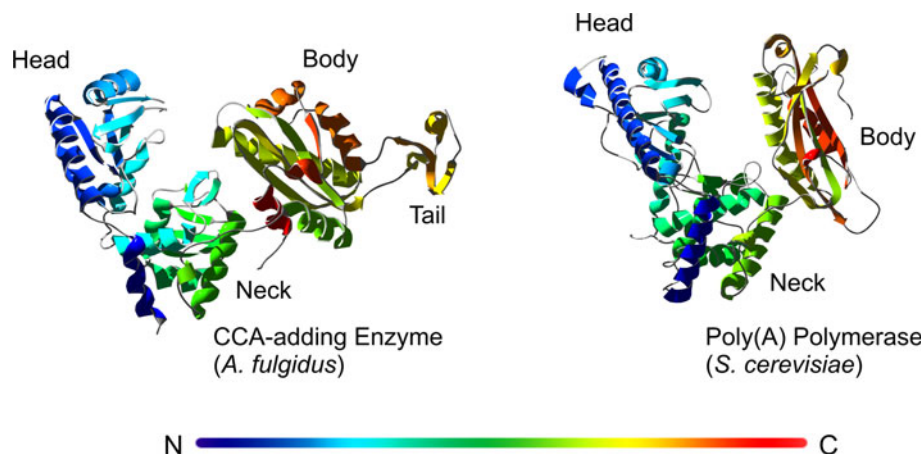
Class I CCA-adding enzymes

General structure

Although catalyzing identical reactions, CCA-adding enzymes found in the three kingdoms of life are remarkably different in the overall structure as well as in the organization of the catalytic core. According to the classification mentioned above, they fall into class I (archaea) and class II (bacteria and eukaryotes) of the polymerase β superfamily. Besides sequence alignments, crystal structure analyses clearly demonstrate the different catalytic strategies these enzymes have evolved. In 2003, the first crystal structure of a class I CCA-adding enzyme, isolated from *Archaeoglobus fulgidus*, was solved [28]. It is organized in four individual domains named *head*, *neck*, *body*, and *tail* in a U-like shape with a cleft between body and tail domains. This architecture has a structural similarity to eukaryotic poly(A) polymerase, confirming the categorization of both enzyme types as class I members (Fig. 1). The *head* domain of the *Archaeoglobus* CCA-adding enzyme corresponds to the nucleotidyltransferase domain of the poly(A) polymerase (PAP), carrying the typical five-stranded β sheet. The *neck* and *body* domains also consist of α-helices and β-sheets and correspond structurally to the central and the RNA-binding domain of PAP [28, 29].

Interaction with the tRNA substrate

While CCA-adding enzymes specifically interact with tRNA molecules (without CCA-terminus) as substrates, it is not the whole tRNA structure per se that is recognized. Shorter minihelix transcripts mimicking the top half of a tRNA, consisting of acceptor stem and T-stem/loop, are also tolerated for CCA-addition [30]. Modeling of enzyme–primer interaction as well as co-crystal structure analyses of a class I CCA-adding enzyme with bound tRNA demonstrated that a cleft in the enzyme architecture has the appropriate size and the required positive electrostatic potential to bind the tRNA top half in the correct orientation for CCA-addition. In this cleft, the tRNA acceptor stem interacts with a highly conserved long α-helical element in an almost parallel orientation [31]. As the position of nucleotide addition, the 3′-end is bound to the active site located in the enzyme's head domain, while the T loop of the tRNA contacts the tail domain [28, 32, 33] (Fig. 5; in comparison with class II enzymes). Interestingly, the enzyme almost exclusively interacts with the sugar–phosphate backbone of the tRNA top half, but not with individual bases or base pairs. Hence, short tRNA duplex structures are recognized as substrates regardless of their nucleotide composition, explaining the enzymes ability to interact with all types of tRNA and tRNA-like molecules without sequence preference [33, 34]. The discrimination against DNA molecules of similar size, however, comes exclusively from interactions with the primer 3′-end, where an aspartic acid residue recognizes the terminal 2′-hydroxyl group of the tRNA molecule [35].

Interestingly, the bound tRNA substrate remains fixed at its binding site in the enzyme during the complete nucleotide incorporation process. This was demonstrated by the finding that a tRNA that was immobilized on the enzyme by crosslinking agents was still a substrate for CCA

Fig. 1 Crystal structures of a class I CCA-adding enzyme (*A. fulgidus*) in comparison with a class I poly(A) polymerase (*S. cerevisiae*). The rainbow color indicates the sequential order of the individual enzyme regions, starting from N- (*blue*) to C-terminus (*red*). Both enzymes show a catalytic cleft that is formed by the head, neck, and body domains. The CCA-adding enzyme carries an additional tail domain. The structures are extracted from the corresponding pdb entries [29, 90]

addition. Furthermore, crystal structures of the enzyme in complex with tRNAs representing the individual reaction intermediates (ending without CCA, with C– and with CC-) indicate that the contacts between tRNA substrate and CCA-adding enzyme do not change during the individual steps of CCA incorporation. These data clearly indicate that the polymerization reaction proceeds without a translocation of the tRNA substrate relative to the enzyme [30, 33, 35]. However, if the tRNA indeed does not move, the growing CCA terminus has to undergo a considerable structural reorganization in the active site of the enzyme in order to make room for the next nucleotide to be added. Such a scenario of a refolding of the tRNA primer after each step of CCA-addition postulates that the enzyme's interaction with the tRNA 3′-end is not very tight and does not contribute to the overall substrate-binding efficiency. This is supported by the fact that mutations around the active site of the *Sulfolobus shibatae* enzyme (also a class I member) that interfere with CCA-addition have only a minor affect on tRNA binding [36]. Hence, it seems that the tRNA body is anchored to the enzyme and ensures an efficient interaction during polymerization, while the tRNA 3′-end as the actual primer site is only loosely associated and is free to refold during nucleotide addition, allowing the re-use of the active site for each individual nucleotide incorporation step.

Catalytic core and reaction mechanism of class I CCA-adding enzymes

Besides binding the tRNA primer, CCA-adding enzymes must also specifically select the nucleotides to be incorporated in a sequence-dependent way. Again, most of the available information for class I comes from the analysis of the *A. fulgidus* enzyme. When crystals of this enzyme were soaked with either ATP, CTP, or UTP, the individual bases were rather disordered and did not show a well-defined electron density. The triphosphate–ribose moiety of these NTPs, however, gave a clear signal. Here, a hydrogen bond between a histidine residue of the enzyme's binding pocket and the 2′-hydroxyl group of the ribose is formed, probably allowing to discriminate against deoxyribonucleotides [28, 29]. While this interaction explains the enzyme's specificity for ribonucleotides, it was quite surprising that it is obviously not able to discriminate efficiently between different nucleotides. However, when the tRNA primer is present, a dramatic increase in the specificity of the single nucleotide binding pocket can be observed. In a ternary complex consisting of *A. fulgidus* enzyme, tRNA and one sort of NTP (either CTP or ATP), the base moiety is positioned in a well-ordered structural arrangement. A highly conserved arginine residue in the NTP-binding pocket of the enzyme and the ribose-phosphate backbone of the 3′-terminal nucleotides of the tRNA acceptor stem form a network of complementary hydrogen bonds with the base (cytosine or adenine) of the incoming nucleotide (Fig. 2). Specifically, N3 of CTP and N1 of ATP interact with the arginine side chain, while the 4-amino group of CTP and the 6-amino group of ATP form hydrogen bonds with the phosphate backbone of the tRNA. This hydrogen bond donor and acceptor pattern is not complementary to the base moieties of UTP and GTP, which are therefore efficiently excluded from binding [33, 37, 38].

For a specific recognition of CTP as well as ATP, the arginine side chain adopts different rotameric positions that allow stable interactions with either CTP or ATP [33, 38]. When CTP is bound, the side chain is oriented towards the bound nucleotide and reduces the size of the binding pocket, leading to a discrimination against ATP. After two additions of CMP, however, the growing tRNA primer is repositioned in the catalytic core and induces a reorientation of the arginine residue, enlarging the binding pocket in order to accommodate ATP. In this specific addition of two CTP and one ATP residue, a β-turn located between
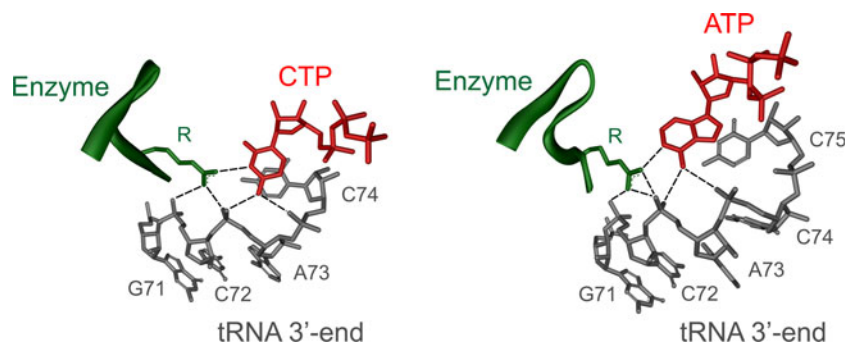

Fig. 2 Nucleotide-binding pocket of a class I CCA-adding enzyme with templating region. A highly conserved arginine residue of the enzyme (*green*) as well as phosphate groups of the tRNA backbone (*grey*) form specific hydrogen bonds with the bound nucleotides (*red*). The collaboration of both protein and tRNA is responsible for an efficient and accurate templating during CCA-addition. The shown structures are parts of the pdb entries of the *A. fulgidus* CCA-adding enzyme [33]

strands 3 and 4 of the nucleotidyltransferase domain plays an important role [36]. Crystal structures as well as mutational analyses indicate that this turn has two essential functions. First, it positions the tRNA 3′-end with the attacking 3′-OH group in the catalytic core for the incorporation of the next nucleotide. Here, it seems that the β-turn is responsible for a 90° flipping of the 3′-terminal base relative to the other acceptor stem positions, leading to an unstacking and repositioning of this terminal nucleotide [33–36]. Correspondingly, the β-turn was suggested to act as a wedge that is positioned between terminal base and the rest of the acceptor stem, preventing stacking interactions between this residue and the tRNA [35]. Second, it is involved in the re-organization of the tRNA part of the NTP-binding pocket so that it can form the required hydrogen bonds with the incoming base. Consequently, the β-turn adopts a different conformation for each step of the polymerization reaction in order to present the 3′-terminal nucleotide of the growing tRNA to the active site [33, 36, 38].

This complex and elegant interplay between enzyme and tRNA primer to select the nucleotides to be incorporated is described best by the term collaborative templating, where both enzyme (protein) as well as substrate (tRNA) form a ribonucleoprotein-based active site that can change its specificity from C to A incorporation [28, 39]. While the enzyme is not moving relative to the tRNA primer, the reorganization of the binding pocket is accompanied by a conformational reorientation of the head domain relative to body and tail regions [38]. After the addition of the last nucleotide, the CCA terminus adopts a continuously stacked conformation and the limited size of the binding pocket does not allow a further repositioning required to flip the terminal A residue for the attack on another NTP. These facts seem to mediate the termination of polymerization after three rounds of nucleotide incorporation, leading to the formation of a tRNA with a defined CCA terminus.

Although the binding of tRNA primer and NTP substrates as well as the sequential addition of two C and one A residue is an obviously highly coordinated and chronological polymerization reaction catalyzed by the CCA-adding enzyme, the basic chemical reaction mechanism is comparably simple and identical for each individual addition reaction. A single active site is catalyzing these three nucleotide additions. It carries two highly conserved carboxylates (DxD) present in the nucleotidyltransferase motif [5] that are essential for catalysis as mutations of either of these residues completely abolish CTP and ATP incorporation. Accordingly, the polymerization reaction takes place by a two metal ion mechanism that was originally described for DNA polymerases [40, 41] (Fig. 3). The two carboxylates bind and position two Mg^{2+} ions in the active
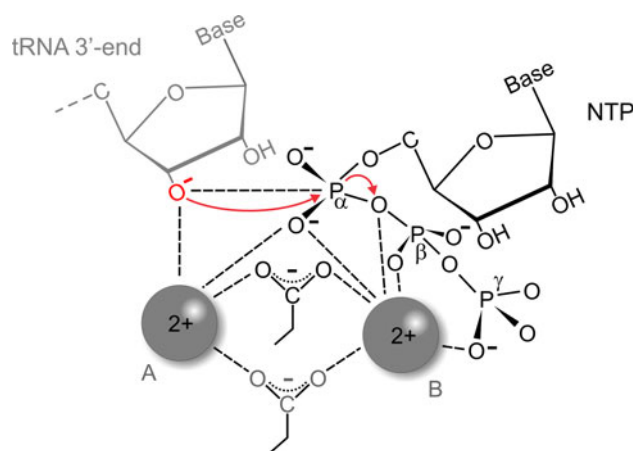

Fig. 3 Polymerization mechanism of CCA-addition according to the general two-metal ion-catalyzed reaction. The two metal ions (Me^{2+}, *grey balls*) are bound to the two catalytically important carboxylates. Metal ion A deprotonates the 3′-OH group of the tRNA primer (*grey*) and activates the resulting 3′-O$^-$ (*red*) for an attack at the α-phosphate of the incoming nucleotide (*red arrows*). The second metal ion B stabilizes the triphosphate moiety of the NTP and facilitates the leaving of the pyrophosphate group. Modified from [40]

site of the enzyme, where metal ion A subtracts the proton from the primer 3′-OH group and facilitates thereby a nucleophilic attack on the α-phosphate group of the incoming NTP. Metal ion B promotes the leaving of pyrophosphate that is released from the bound NTP. In addition, both metal ions stabilize the transition state of this reaction. These events are repeated for each individual step of CCA-addition in the catalytic core of the enzyme.

Quality control during CCA-addition

A puzzling fact concerning CCA-addition is the extreme fidelity of this reaction, given that the CCA-adding enzymes are independent of a nucleic acid template and show no proofreading mechanism. The collaborative templating of enzyme and tRNA in combination with a β-turn that positions the 3′-end of the primer and helps to reorganize the binding pocket obviously guarantee a highly accurate nucleotide addition. However, there seems to be another checkpoint during CCA-addition that contributes to the fidelity of the reaction [37]. When the *A. fulgidus* enzyme was tested for A addition to RNA substrates ending with -NC or -CN (N represents A, U or G residues instead of C74 and C75 of the CCA sequence), the incorporation of the terminal A residue was either completely abolished (on primer ends carrying A or G residues) or dramatically reduced (on primer ends with U residues). CTP incorporation, on the other hand, was more promiscuous and accepted primers with altered position 74. Hence, at least for the addition of the terminal A residue, there exists a quality control for the tRNA 3′-end that

prevents the formation of aberrant and non-functional CCA termini. This interpretation is corroborated by crystal structures of the corresponding complexes consisting of enzyme and tRNA minihelices. For substrates ending with N74, all complexes showed an open conformation like that of the complex with the correct substrate (ending with C74). Correspondingly, these substrates were accepted for addition of CMP at position 75. Substrates ending with -NC or -CN that were not accepted for A incorporation, however, did not lead to enzyme complexes in a catalytically active conformation.

Together with the intrinsic fidelity of the CCA addition, this quality checkpoint leads to the amazing accuracy of these enzymes. Furthermore, if one assumes that tRNAs with erroneous 3′-ends are degraded in the cell, the amount of such incorrect tRNAs will never reach a level deleterious to the cell [37].

Class II CCA-adding enzymes

In principle, class II CCA-adding enzymes catalyze the same type of reaction as class I enzymes do. Yet, this group differs dramatically from its class I counterpart not only in primary sequence and overall architecture but also in the catalytic strategy for a template-independent nucleotide incorporation at a similar fidelity [5, 34, 42, 43].

Overall structural architecture

The crystal structures of class II CCA-adding enzymes from *B. stearothermophilus*, *H. sapiens* and *T. thermophilus* have been solved in the apo form as well as in combination with bound nucleotides or tRNA primer [44–46]. These enzymes share a high structural homology with a hook- or seahorse-like shape composed of head, body, neck and tail domains [45] (Fig. 4). In contrast to class I enzymes, only the head domain carries an antiparallel β-sheet, while neck, body and tail consist exclusively of α-helical elements. The head domain is homologous to the palm domain of DNA polymerase β and contains the highly conserved carboxylates that coordinate the catalytically essential metal ions [44, 45]. In contrast to polymerase β, these enzymes lack a thumb domain that is required for binding a double-stranded DNA primer [45].

Compared to class I, class II CCA-adding enzymes show a much higher evolutionary conservation of individual catalytic core motifs. The active site is located in the N-terminal part of the enzyme and consists of five elements (motifs A to E) that are involved in metal ion binding, catalysis, ribose recognition, nucleotide selection, and templating [32, 44–46].

Motif A is located in the head domain and includes the general signature motif of all nucleotidyltransferases with the two metal-binding carboxylates DxD (x is any amino

Fig. 4 Structural organization of class II CCA-adding enzymes. *Upper part* The catalytically important motifs are indicated as *red boxes* with consensus sequences in the rainbow bar. In the three-dimensional structure (CCA-adding enzyme from *Thermotoga marimita* [46]), these motifs are indicated in *dark red* in the head and neck domain. *Lower part* The nucleotide-binding site (*green*) recognizes the incoming CTP and ATP (*red*) by the formation of Watson/Crick-like hydrogen bonds between the amino acid template EDxxR and the corresponding edge of the bases. The nucleotide-binding site structure is derived from the pdb entry of the *B. stearothermophilus* CCA-adding enzyme [45]
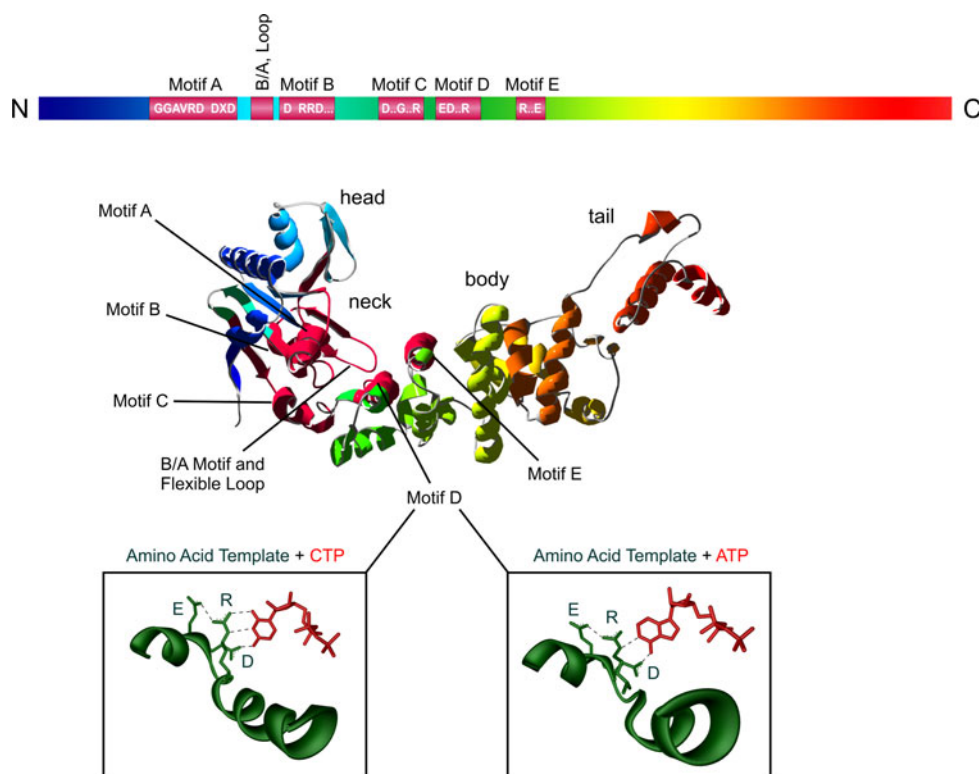
acid) that are involved in catalysis and binding of the triphosphate moiety of the incoming nucleotides. Additionally, the head domain carries motif B, where highly conserved residues play a critical role in discriminating between NTPs and dNTPs. The neck domain contains motif D, a single nucleotide-binding pocket that is specific for binding of CTP and ATP [45]. Furthermore, head and neck domain form a cleft that binds the incoming nucleotide as well as the 3′-end of the tRNA primer. The body and tail domains at the enzyme's C-terminus recognize the top-half region of the tRNA primer [32, 46].

Similar to class I, the class II CCA-adding enzymes also have a poly(A) polymerase (PAP) counterpart. Here, structural similarities could not be identified yet, as crystal structures of class II PAP are still missing. At the sequence level, however, the poly(A) polymerases show a surprisingly high homology with all conserved elements of the catalytic core described above, while only very few sequence motifs specific for poly(A) polymerases could be identified [47].

CCA-adding enzymes recognize their tRNA substrates via size, shape, and charge complementary

Similar to class I enzymes, class II CCA-adding enzymes recognize predominantly the top half of a tRNA molecule. In the crystal structures of several CCA-adding enzymes, a positively charged cleft between head and neck domains was observed and led to the assumption that the tRNA substrate binds to this region [32, 45, 46]. This idea was confirmed by crystal structures of enzyme in complex with tRNA, where the tRNA top half is located exactly in this cleft [32]. In addition, the 3′-end of the tRNA primer is positioned close to the active site and forms hydrogen

bonds to several conserved amino acid residues. Correspondingly, tRNA minihelices consisting of acceptor stem and T-arm are readily accepted by the enzyme for CCA-addition [20, 48, 49]. Furthermore, it seems that—like in class I—the CCA-enzyme does not move along the tRNA during synthesis but remains at a fixed position [30]. This interaction of tRNA and enzyme indicates that the growing 3′-end must refold in the active site in order to accommodate the primer for further nucleotide incorporation (see below). Obviously, the tRNA binding without slippage restricts the enzyme to catalyze only one round of CCA-addition. This anchoring function was shown in chimeric enzymes where the C-terminal part of the CCA-adding enzyme was replaced by the corresponding region of a related poly(A) polymerase (see below: closely related class II nucleotidyltransferases). The resulting chimeras carried the complete catalytic core of the CCA-adding enzyme and synthesized accordingly the CCA triplet on tRNA primers. However, instead of stopping after one CCA completion, the chimeras added multiple rounds of CCA triplets, leading to a poly(CCA) tail, comparable to polyadenylation. Hence, the original C-terminus of the CCA-adding enzyme obviously inhibits this further polymerization, and it was speculated that the tRNA substrate is bound in a fixed position that does not allow any gliding of the enzyme along its substrate as believed to take place in poly(A) polymerases [50].

Interestingly, the tRNA 3′-end enters the catalytic core of class II CCA-adding enzymes from the opposite direction as in class I enzymes (Fig. 5). However, as both classes use the same common mechanism of nucleotide transfer, the 3′-terminal nucleotide has to rotate in class II enzyme in order to adopt a position compatible with a nucleophilic attack on the next NTP to be incorporated.
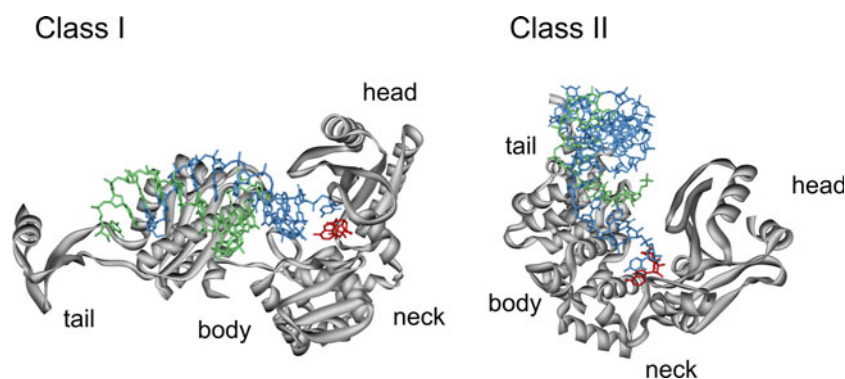

Fig. 5 Binding of the tRNA primer in class I and II CCA-adding enzymes. For a better visibility of tRNA and bound ATP, the orientation of the enzymes was reversed compared to Figs. 1 and 4. In order to compare the entry sites of the tRNA primer (*blue*, strand with 3′-OH, *green*, complementary strand with 5′-end), the bound ATP (*red*) is oriented in identical positions in the proteins (*grey*). While in class I enzymes (*left*), the primer enters the enzyme from the left, it is

bound to class II from the top. Due to these dramatically different orientations, the 3′-terminal base of the primer in the class II enzyme has to leave its original position in the tRNA and has to rotate for about 90° in order to be in a position that allows a stacking interaction with the bound NTP and an efficient nucleotide transfer. Structures are extracted from the corresponding pdb entries [32, 33]

Consequently, after this rotation, the primer 3′-termini have very similar positions in the catalytic cores of both enzyme classes [32, 34, 43] (Fig. 5).

Apart from the few specific interactions between the 3′-terminal bases of the tRNA and the catalytic site, crystal structures indicate that the enzyme's body and tail domains interact almost exclusively with the sugar–phosphate backbone of the tRNA top half, whereas the anticodon region protrudes from the enzyme [32, 33, 46]. Hence, class II enzymes use exactly the same strategy as class I to recognize a whole set of tRNA molecules with different sequence composition. Furthermore, this interaction of the enzyme with the T-loop might represent the anchor function that avoids a sliding of the enzyme along the tRNA substrate, allowing only one round of CCA-addition [32, 50].

Taken together, CCA-adding enzymes of class I and II use size, shape and charge complementary in order to recognize their tRNA substrates and to distinguish them from other RNAs present in the cell.

Class II enzymes carry a set of highly conserved motifs important for catalysis

While in class I enzymes the nucleotidyltransferase signature is the only highly conserved sequence element, class II enzymes carry a whole set of equally conserved motifs with defined individual roles in catalysis. These motifs are located in the N-terminal part of the enzyme and come into rather close contact in the three-dimensional structure in order to form the active site as well as the nucleotide-binding pocket [32, 44–46] (Fig. 4).

The most N-terminally located element is motif A, which corresponds to the nucleotidyltransferase motif. It carries the described two catalytically active carboxylates DxD that position the two metal ions required for the common polymerization process. As this mechanism is already described for class I enzymes (see above), it is not further explained at this point. In the second element (motif B), located in the head domain, the highly conserved sequence RRD is found that recognizes the 2′-OH group of the bound NTP and discriminates against deoxynucleotides [45]. The guanidinium group of the central arginine residue of this amino acid triplet forms a single hydrogen bond with the 2′-hydroxyl group of the incoming nucleotide. When this arginine side chain is replaced by isoleucine, the resulting enzyme variant does no longer discriminate against deoxynucleotides and readily adds a dCdCdA sequence to the tRNA primer [51].

While no specific function is assigned to motif C, motif D (located in the neck domain) plays an important and fascinating role during CCA-addition. Although biochemical analyses pinned down a single nucleotide binding site

for the three individual NTPs to be incorporated, it remained unclear how the enzyme specifically recognizes CTP and ATP and discriminates against the remaining two nucleotides [30, 39]. This question was answered by the structural analysis of enzyme crystals soaked with ATP and CTP [45]. A set of three conserved amino acid positions EDxxR (x represents any amino acid) in motif D was shown to form Watson–Crick-like hydrogen bonds to the base moieties of the bound nucleotides (Fig. 4). Hence, these amino acids have a true templating function comparable to that of nucleic acid templates in DNA replication or transcription. For CTP binding, the guanidinium group of the arginine residue forms two hydrogen bonds with the N3 and O2 positions of the Watson–Crick interface. A third bond is contributed by the aspartic acid side chain, recognizing the 4-amino group of the base. As a result, enzyme and base form the same set of hydrogen bonds that is found in base-paired nucleic acids. While glutamic acid as the third amino acid residue does not interact directly with the bases, it plays an important role in the reorganization of the binding pocket in order to switch its specificity from CTP towards ATP binding. In this rearrangement, the glutamate together with the arginine residue rotates and enlarges the binding pocket to accommodate the larger adenine base. Subsequently, two hydrogen bonds (typical for A–T or A–U base pairs) are formed. The arginine interacts with the N1 position of ATP, while the aspartic acid residue forms a hydrogen bond with the 6-amino group. While these hydrogen interactions are highly similar to the ones found in nucleic acid base pairs, the possibility to switch the base-pairing properties is a unique and fascinating feature of this amino acid template.

Furthermore, the selection mechanism of this protein-based template explains why UTP and GTP are efficiently excluded from binding to the enzyme's active site, as these bases show a different and therefore incompatible pattern of hydrogen bond donors and acceptors [45]. Consequently, an adjustment of the hydrogen bond interface in the protein to interact with UTP and GTP should lead to a switch in NTP specificity of the enzyme. Indeed, when the EDxxR template was replaced by NNxxE or SNxxE, the resulting binding pockets could form the corresponding hydrogen bonds with guanine and uracil, leading to variants that favored GTP and UTP over CTP and ATP [51].

The importance of this amino acid template for nucleotide selection was further demonstrated by replacing the arginine position by alanine which cannot form the required hydrogen bonds with CTP and ATP. Consequently, the resulting enzyme accepted any nucleotide for incorporation, demonstrating a dramatic loss in specificity. A replacement of the other two templating residues (glutamate and aspartate), on the other hand, had only minor

effects on the fidelity of the enzyme [20, 51]. However, when all four NTPs were offered simultaneously, the enzyme variant with the mentioned R to A replacement showed an astonishing fidelity in CCA-addition, leading to a considerable amount of tRNAs with correct CCA ends, while tRNAs with misincorporations were not further elongated [20]. Obviously, the CCA-adding enzyme controls for the incorporation of correct nucleotides and stops polymerization after an incorrect nucleotide was added. Although the mechanism underlying this quality control is not fully understood yet, it is discussed that the last incorporated nucleotide is inspected and identified by a further set of amino acids forming hydrogen bonds to the base moiety. Two of these amino acids are located between motif A and B and form a small but highly conserved element consisting of a basic residue followed by any amino acid and a subsequent acidic side chain, described as basic/acidic motif (B/A) [32, 46, 52] (Fig. 4). As indicated by crystal structures as well as mutational analyses, the basic residue (in most cases an arginine) forms a hydrogen bond to the O2 position of the tRNA's C75. This and other putative interactions seem to bring the 3'-OH of the tRNA end into position for a nucleophilic attack at the α-phosphate of the next incoming nucleotide. If a nucleotide other than CTP was incorporated, the 3'-end cannot be positioned for the next nucleotide addition. Hence, only tRNAs with correctly added residues are a proper substrate for further incorporations. These findings show that class II enzymes also possess a quality checkpoint comparable to class I enzymes. However, class II nucleotidyltransferases obviously can use this mechanism as a back-up system that allows a proper CCA-addition even in the presence of a rather deleterious mutation in the amino acid template [20].

Although in class II enzymes the tRNA primer does not play a templating role as observed in class I, the back-up mechanism indicates that RNA and enzyme also interact and collaborate in CCA-addition. Structures of ternary complexes of class II enzymes underline the existence of such a dynamic ribonucleoprotein architecture [43, 46]. Whereas RNA and protein form a base-recognizing template in class I enzymes, the protein part alone is responsible for nucleotide selection in class II. Here, the tRNA fulfils a different function in controlling the fidelity of the incorporated nucleotide. Hence, two different and surprising solutions for the same task (an error-free synthesis of CCA-termini) evolved for the two types of CCA-adding enzymes.

The function of the last motif E is not fully understood yet. It seems to stabilize a helix-turn structure found in motif D and might also bind to the tRNA primer [45]. However, besides these five motifs A to E, the catalytic core of class II enzymes contains another element that is involved in the specificity switch of the nucleotide-binding

pocket and that differs dramatically from the above-mentioned elements in terms of function and evolution.

During CCA-addition, the shift from CTP to ATP binding is achieved by the reorientation of the templating amino acids in motif D as described above. However, this reorganization of the NTP-binding site is not a locally restricted rearrangement of an individual protein region, but seems to be involved in an extensive movement of individual enzyme domains relative to each other. This is supported by the fact that a co-crystal of a class II nucleotidyltransferase with bound tRNA-CC readily dissolved upon soaking with ATP [32]. Furthermore, a kinetic analysis of the E. coli CCA-adding enzyme showed that an increased concentration of glycerol (35%) in the reaction buffer leads to a considerable reduction of A incorporation. This result suggests that the enzyme has to perform an extensive movement for this reaction that is impeded by the high viscosity of glycerol [53]. A prerequisite for such a structural rearrangement is the existence of flexible elements that allow a movement of individual protein parts relative to each other. A corresponding region was identified in the crystal structures of several class II enzymes. Due to a highly disordered conformation, a loop of 10 to 20 amino acids located in the catalytic head domain was not resolved in most of the structures, indicating that this protein part is highly flexible [32, 44–46, 52].

Direct proof that this loop element is involved in A incorporation was demonstrated by the introduction of several mutations. Deletions as well as single point mutations completely abolished the A addition, while the incorporation of the two C residues remained unaffected [32, 46, 52, 54–56]. Since the loop sequence of the E. coli enzyme has some similarities to P-loop elements, a possible function in ATP binding was discussed [56]. However, a loop mutation that strongly interferes with ATP incorporation still has an unchanged K_M, indicating that ATP binding is not affected by this mutation. Consequently, a direct and specific interaction with ATP can be excluded [55]. Hence, the function of the flexible loop seems to be rather indirect, acting as a hinge or lever that allows a movement of the adjacent enzyme parts leading to the required conformational reorganization of the nucleotide-binding pocket [52]. Due to such an important function, one would expect that this element shows a similarly high conservation at the sequence level as the other motifs of the catalytic core do. However, phylogenetic analyses revealed that this is not the case, as the loop sequences of individual CCA-adding enzymes show a remarkable sequence diversity [52, 54] (Hoffmeier, unpublished data). Nevertheless, when the loop sequences were assigned to the organisms the analyzed CCA-adding enzymes were isolated from, individual consensus sequences could be recognized that allowed to classify the loops into distinct evolutionary

conserved families. Interestingly, these sequence families turned out to be rather incompatible, as reciprocal loop exchanges completely abolished the A-adding activity in the resulting chimeric enzymes, while exchanges within these families were tolerated. This indicates that the loop elements are only functional within a certain enzyme context of the same phylogenetic origin, probably interacting with amino acid residues in the enzyme's body. Consequently, it is very likely that the loop is not functioning as a passive hinge module, but serves as a lever that pulls the amino acid template of the nucleotide binding pocket into the required position. One of these possible contact points was identified in a recently solved crystal structure of the CCA-adding enzyme of *Thermotoga maritima*. Here, a tyrosine residue of the loop interacts with an aspartate side chain at the first position of the amino acid template (in *T. maritima*, the templating motif carries the sequence DD*xx*R, a slight deviation from the usual ED*xx*R motif). This interaction might influence the orientation of the templating arginine residue and contribute to the specificity switch [46]. However, as this tyrosine residue is not conserved in all loop families, this interaction cannot be generalized. Rather, it seems more likely that other positions within the different loop subtypes can interact with the templating motif. It is also conceivable that residues located outside of the amino acid template are recognized and form the lever structure. The latter scenario implies that these binding partners underlie a co-evolution together with the individual loop sequences (Hoffmeier, unpublished data). As a consequence, the loop does not follow the usually observed co-evolution of catalytically important motifs in order to optimize the enzymatic activity, but has its own evolutionary path that depends on sequence elements of the individual CCA-adding enzymes that are not identified yet.

While these elements involved in CCA-addition are found exclusively in class II enzymes, it seems that both classes carry a important β-turn with identical function in the catalytic core. In class I, the turn has a conserved sequence and is located between β-strands 3 and 4 of the nucleotidyltransferase domain, where it is involved in positioning the tRNA 3'-end for proper nucleotide incorporation (see above). In class II, a comparable turn is found between strands 4 and 5. While the sequence of this turn is only moderately conserved, the structural organization is very similar to corresponding turns in other nucleotidyltransferases, as revealed by structural overlays [46, 47]. Accordingly, the function of this turn might be identical between class I and II CCA-adding enzymes.

Taken together, these elements identified in class II enzymes build up a complex three-dimensional structure that brings the two substrates tRNA and nucleotides into close proximity required for an efficient and accurate CCA-addition. The tRNA substrate, on the other hand, serves not only as a primer to accept further nucleotides, but has a significant contribution to the fidelity of the reaction. Although it is not directly involved in NTP recognition as in class I, it is a similarly active component of the class II ribonucleocomplex. As these enzymes also carry just a single nucleotide-binding site, the growing 3'-end of the tRNA must progressively refold and allow the binding of the next NTP to be incorporated. This refolding occurs as a concerted movement together with the enzyme, triggering the specificity switch to ATP and—together with the anchor function of the C-terminal domain—probably also the termination of the reaction.

Alternative substrates for CCA-adding enzymes

Just like the class I enzymes, the CCA-adding enzymes of class II recognize only the upper part of a tRNA molecule (consisting of acceptor stem and T-arm) as a substrate, but not the complete tRNA structure (Fig. 5). Hence, it is not surprising that for both classes, artificial minihelices can be used for efficient CCA-addition in vitro [20, 48, 49]. However, minihelix-like structures at the 3'-end of transcripts are also present in vivo. In eukaryotic cells harboring class II CCA-adding enzymes, there are many cases where such structures carry additional non-encoded C and A residues at the 3'-ends, probably added by class II CCA-adding enzymes. One of the first examples was found in tobacco mosaic virus, where individual transcripts carried complete CCA termini [57]. Similarly, considerable amounts of U2 snRNA, found in the nucleus of human cells, end with a CCA sequence, where at least the terminal A residue was incorporated posttranscriptionally [58]. Further examples for CCA ends at non-tRNA transcripts were described in maize mitochondria, where *rps12*, *cox2* and *atp9* mRNAs have non-encoded 3'-terminal C and A residues [59], as well as in *Arabidopsis* [60] and in chloroplasts of tobacco [61]. Such alternative in vivo substrates for CCA-addition can certainly be found for class I enzymes as well. Yet, this type of CCA-adding enzymes exists only in archaea, where non-tRNA-like molecules were not analyzed yet in this aspect. However, the biological relevance of these events of CCA-addition to non-tRNA like transcripts is unclear. It is rather likely that these RNA molecules meet the substrate requirements of a CCA-adding enzyme just by chance.

CCA-addition is stimulated by Hfq

As described above, class II enzymes catalyze the CCA-addition at high efficiency and accuracy in vitro, without the help of any assisting proteins or co-factors. In vivo, however, the situation might be different, at least in some

cases. For the *E. coli* enzyme, it was shown that the multifunctional protein Hfq, originally discovered as a host factor for phage Qβ [62], can stimulate the CCA-adding activity [63]. Whereas a direct interaction with the enzyme could not be documented, Hfq binds specifically to tRNAs [63, 64]. A kinetic analysis of the stimulated reaction revealed that the affinity of the CCA-adding enzyme to the tRNA/Hfq complex was not increased compared to the naked tRNA alone. However, it is very likely that Hfq facilitates the release of the reaction product, after CCA-addition took place, as kinetic analyses suggest that the enzyme remains rather tightly bound to the completed tRNA, representing the rate-limiting step of the reaction [53, 63].

While the described Hfq effect is currently the only documented case of a stimulation of CCA-addition, it is possible that similar proteins also affect tRNA maturation and maybe also CCA-addition. Hfq belongs to the Sm/Lsm-like protein family, and the corresponding proteins in yeast have already been demonstrated to play an important role in the processing of several tRNA transcripts [65]. Furthermore, enzymes related to class II CCA-adding enzymes are also likely to be stimulated by Hfq. For the structurally closely related bacterial poly(A) polymerase, such a stimulating effect has already been demonstrated [66, 67].

Some class II enzymes have an additional function as phosphatases

In contrast to the highly conserved N-terminal region containing the catalytic core, the C-terminus of class II CCA-adding enzymes shows a remarkable sequence variation and is obviously not required for CCA-addition, as considerable parts can be deleted without affecting the enzymatic activity [50]. However, in some bacteria like *E. coli*, the C-terminus of the CCA-adding enzyme carries an HD domain characteristic for metal dependent phosphohydrolases [68, 69]. Consequently, these enzymes have an efficient phosphatase activity and can remove 3′-phosphate groups from tRNA substrates [20]. While in the first instance, this additional moonlighting activity is surprising, it definitely represents a very useful function in the cell. RNA molecules are generally rather susceptible to hydrolytic damage, induced by spontaneous cleavage or nucleolytic activity of RNases. This cleavage leads to 2′–3′ cyclic phosphate groups at the 3′-end of the transcripts [70, 71]. A corresponding damage at the tRNA CCA-terminus results in tRNAs with shortened 3′-ends carrying cyclic phosphate groups which are not functional anymore. Furthermore, the cyclic phosphate group blocks the 3′-hydroxyl for a subsequent CCA-end restoration. The phosphatase activity of the CCA-adding enzyme, however,

can remove this cyclic phosphate group, regenerating the 3′-OH group that is then used as a substrate for the incorporation of the missing parts of the CCA-end [20]. Hence, in *E. coli*, the CCA-adding enzyme acts as a tRNA repair enzyme with a coordinated and sequential activity as a phosphatase and a nucleotidyltransferase (Fig. 6). Since this described damage of tRNA molecules is ubiquitous and occurs therefore also in organisms lacking such a phosphatase/nucleotidyltransferase combination, it is very likely that CCA-adding enzymes without HD domains collaborate with separate phosphatases in order to provide an efficient tRNA 3′-end repair.

Closely related class II nucleotidyltransferases

While each organism usually carries just one single gene for a tRNA nucleotidyltransferase, there are several bacterial species which obviously express two such enzymes. Interestingly, these enzymes do not synthesize the complete CCA-end, but show a restricted activity. One enzyme adds exclusively two C residues to the tRNA primer, while the second one incorporates the terminal A. Consequently, these enzymes are called CC- and A-adding enzymes, respectively [52, 72–75]. The collaboration of both activities ensures a complete and efficient CCA synthesis. A similar CC-adding enzyme was also found in humans as a result of an alternative splicing event that removes two α-helices located in the body of the enzyme. However, it is not understood yet why a deletion in this region restricts the activity to the addition of two C residues. Furthermore, a corresponding A-adding enzyme has also not yet been identified in humans. Hence, while this splicing event seems to be conserved in eutherian mammals, its biological role in these organisms is still unclear [76].

As both CC- as well as A-adding enzymes carry the complete set of conserved motifs A to E, it was rather difficult to distinguish between the individual activities at the sequence level. In a detailed analysis based on sequence alignments, a common deletion in CC-adding enzymes could be identified. In these enzymes, the flexible loop required for the specificity switch from C to A incorporation is missing. Without this lever element, the nucleotide-binding pocket of these proteins is fixed in a conformation that allows only the interaction with CTP, while the structural reorganization required for ATP recognition is no longer possible. Experimental evidence comes from an analysis of closely related *Bacillus* enzymes, where the loop of a CCA-adding enzyme was transplanted at the corresponding position of a CC-adding enzyme. The resulting chimera showed a full CCA-adding activity, representing therefore a gain of function due to the loop insertion. Accordingly, the artificial deletion of this loop completely abolished the A incorporation of the CCA-
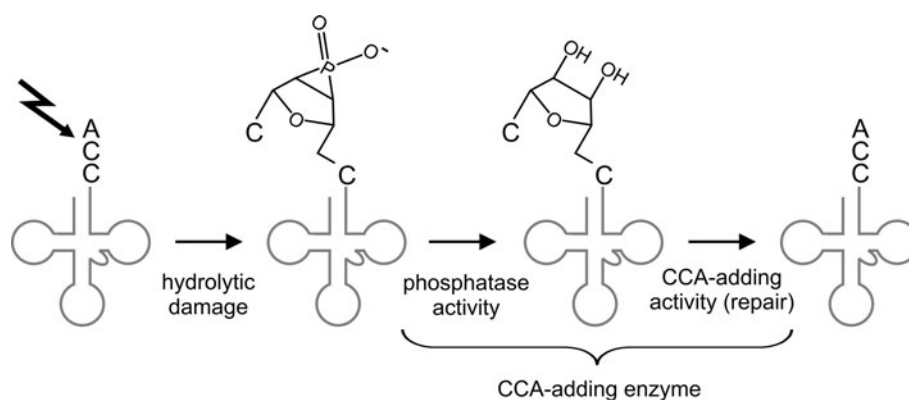
Fig. 6 Repair scenario catalyzed by a class II CCA-adding enzyme with phosphatase activity. Due to hydrolytic damage (*black bolt*), the tRNA (*grey*) has lost the terminal A residue of the CCA-end and carries now a 2′3′-cyclo phosphate. The phosphatase center of the CCA-adding enzyme removes this phosphate group and converts the tRNA 3′-end into a standard primer, carrying 2′- and 3′-OH groups. Subsequently, the nucleotidyltransferase activity of the CCA-adding enzyme incorporates a new A residue and restores the CCA terminus

adding enzyme from *E. coli* [52]. However, when the loop from the *Thermotoga maritima* CCA-adding enzyme was inserted into the *Aquifex aeolicus* CC-adding enzyme, the resulting chimera showed no complete CCA-addition [46]. Obviously, the *Thermotoga* loop is not functional in the *Aquifex* enzyme body, supporting the incompatibility of the loop families described above. Nevertheless, the existence of this deletion can be used as a hallmark to identify CC-adding enzymes. For the A-adding enzymes, on the other hand, the molecular basis for the restriction of their catalytic activity to the addition of the terminal A residue is still not understood. Neither deletions nor mutations in the catalytically important motifs have been identified yet. Consequently, the enzymatic activities have to be determined experimentally. In a phylogenetic analysis, however, the A-adding enzymes are found in one separate branch of the tree, while CC-adding enzymes and CCA-adding enzymes do not strictly separate [52]. Hence, it is conceivable that the A-adding enzymes are of monophyletic origin and evolved in an early separating individual lineage.

Another closely related member of class II CCA-adding enzymes is the bacterial poly(A) polymerase (PAP). This enzyme, although not present in all bacterial species [6, 47], has an important function in RNA quality control. In the bacterial system, the poly(A) tails do not confer mRNA stability, but represent signals for exonucleolytic degradation by the RNA degradosome [77–79]. Unlike CCA-adding enzymes, no crystal structures of bacterial poly(A) polymerases have yet been determined. Consequently, the exact mechanism of polyadenylation and the differences to CCA-addition are not fully understood. However, sequence alignments indicate a very close evolutionary relation to the class II CCA-adding enzymes. Especially the N-terminal catalytic core is highly similar and carries the same set of catalytically important motifs [50, 80, 81]. Based on this high sequence similarity, it has been speculated that CCA-adding enzymes and poly(A) polymerases share a common ancestor and might have interconverted during evolution [5]. Indeed, the conversion of a poly(A) polymerase into a CCA-adding enzyme was demonstrated by the generation of a chimera, where the C-terminal part of the *E. coli* PAP was replaced by the corresponding region of the CCA-adding enzyme from the same organism [50]. The resulting chimeric enzyme did not produce poly(A) tails, as would have been expected due to the catalytic core of PAP, but exclusively synthesized CCA-termini on tRNA substrates. This surprising CCA-addition is explained by the fact that the nucleotide-binding pocket of the poly(A) polymerase carries an amino acid template ED*xx*R identical to that found in the CCA-adding enzyme. Hence, the binding pockets of both enzymes have the intrinsic ability to recognize ATP as well as CTP. However, while the back-up mechanism described for CCA-adding enzymes compensates for mutations in this amino acid template (see above), poly(A) polymerases with a mutated binding pocket show a dramatically increased misincorporation [54]. Obviously, the poly(A) polymerases do not have a back-up mechanism for polyadenylation and the polymerization specificity does not rely on the collaboration between RNA substrate and enzyme, as it was observed for the CCA-adding enzymes. Whereas this quality control checkpoint in the CCA-adding enzymes underlines the vital importance of a correct CCA-addition in the cell, poly(A) tails with misincorporated additional nucleotides are tolerated to a certain extent, making a similar back-up system dispensable [54, 82, 83].

Besides the amino acid template required for nucleotide selection, bacterial poly(A) polymerases also carry the flexible loop lever region originally identified in the CCA-adding enzymes. However, while this element is essential in CCA-adding enzymes for the incorporation of the terminal A residue, it is not required for polyadenylation, as
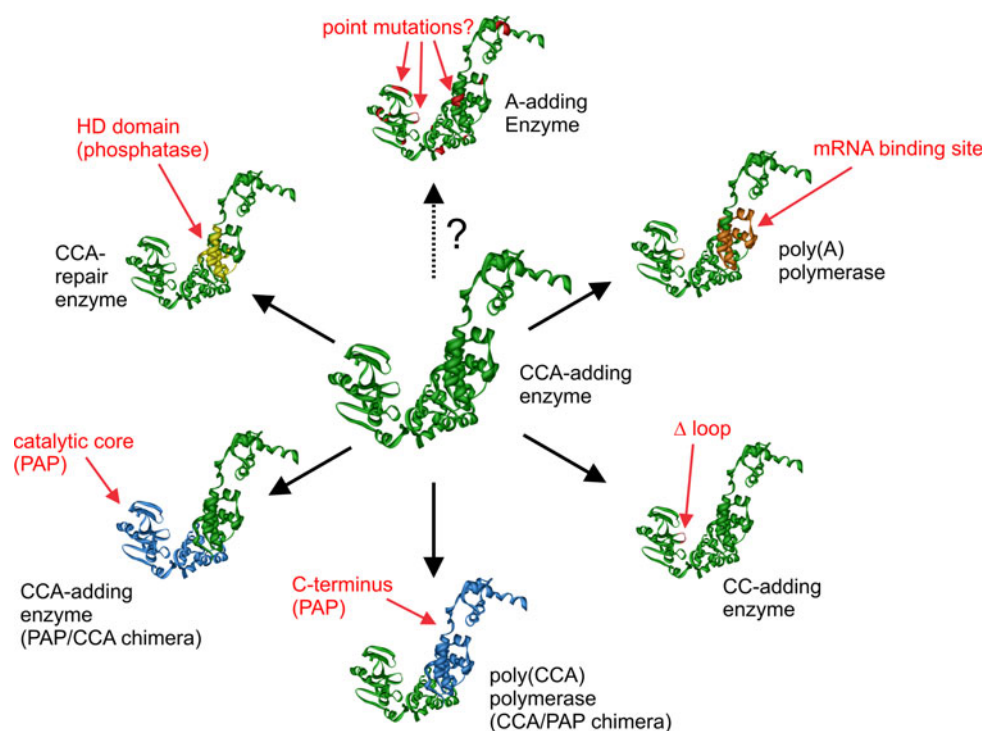
Fig. 7 Modular evolution of class II nucleotidyltransferases as supported by experimental and phylogenetic data. From a starting CCA-adding enzyme (*center*), new enzymes emerge by the insertion/deletion/exchange of protein modules. Individual events of this scenario are presented in a clockwise orientation: An mRNA-binding site module transforms the enzyme into a poly(A) polymerase. The deletion of the lever loop module restricts the activity to CC-addition.

Exchange of large N- or C-terminal parts lead either to a poly(CCA) polymerase or to a further CCA-adding enzyme with the catalytic core of poly(A) polymerase. The insertion of the HD domain introduces a phosphatase activity. The evolution of A-adding enzymes, however, is not yet clarified (*dashed arrow with question mark*). The individual insertions, deletions, and replacements are indicated only at approximate positions

the corresponding deletion variant of PAP readily synthesizes long stretches of poly(A) tails [54]. These data clearly support the idea that—instead of an evolutionary inter-conversion—class II poly(A) polymerases descend from CCA-adding enzymes and still carry active site elements important for CCA-addition as an evolutionary relic.

Evolution of class II nucleotidyltransferases

The high sequence similarity of CCA-, CC-, and A-adding enzymes as well as the bacterial poly(A) polymerases led to an ongoing debate concerning their evolution. Based on phylogenetic analyses it was hypothesized that CC- and A-adding enzymes might represent the ancestral rather primitive enzymatic activity that evolved into the more complex CCA-adding enzymes that can switch from CTP towards ATP specificity [74, 75]. In the case of the A-adding enzymes, this idea is supported by two facts. First, the *Thermotoga maritima* CCA-adding enzyme has a striking similarity to A-adding enzymes, indicating that it evolved out of this type of nucleotidyltransferase [75]. Second, phylogenetic analyses indicate that the A-adding enzymes fall together in a separate branch of the

evolutionary tree, from which CC- and CCA-adding enzymes might have emanated [52]. On the other hand, the same phylogeny indicates that the CC-adding enzymes descend from CCA-adding enzymes, since some of them are located in the same branch as the CCA-adding enzymes. This event obviously occurred at least twice in evolution, as CC-adding enzymes are found in two different branches of the tree [52]. This scenario is corroborated by the observation that the deletion of the lever loop is responsible for the reduced activity of these enzymes. It is much more likely that such a deletion occurs during evolution than an insertion of a loop element into a CC-adding enzyme at the correct position. As there are many incompatible loop families, this insertion has to have a sequence composition that is tolerated by the catalytic core region of the CC-adding enzyme, which is also highly unlikely. Further support for the idea that a CCA-adding enzyme evolved into CC-adding enzymes comes from kinetic analyses [51]. Hence, at least for these two types of enzymes, the evolutionary path seems to be clarified. In addition, it is also very likely that the closely related bacterial poly(A) polymerase directly descends from the CCA-adding activity, as it still carries

elements required for CCA-addition but unnecessary for polyadenylation. Taken together, the available experimental as well as phylogenetic data indicate that the CCA-adding activity represents the ancestral state that evolved into different further activities based on a combinatorial exchange of individual functional modules. Thus, the observed change in activity due to this transfer of enzyme parts represents an impressive example of the modular composition of proteins that is frequently discussed as one of the driving forces in enzyme evolution [84–86] (Fig. 7).

Compared to class II, the class I nucleotidyltransferases represent a much bigger group with members catalyzing a broad variety of reactions [6]. Yet, no tRNA nucleotidyltransferases with partial activities or enzymatic interconversions of CCA-adding enzymes and poly(A) polymerases were described up to now. Hence, these interesting evolutionary events obviously represent a specific feature of class II enzymes. Nevertheless, it is very likely that class I enzymes also evolved out of an ancestral primitive nucleotide transferring activity similar to the MNTs mentioned in the introduction. By the addition of new functional domains, these MNTs evolved into the two separate classes I and II with different CCA-adding enzymes as well as poly(A) polymerases. Then, further combinatorial module exchanges led to the observed variety of enzymatic activities within class II (Fig. 8).

CCA-adding enzymes: molecular fossils of an ancient telomerase?

Although modern class I and class II CCA-adding enzymes use different strategies to synthesize the CCA sequence on tRNA 3′-ends at a surprisingly high fidelity, the shared ability to accept minihelical structures as primers might indicate a common evolutionary origin. The top half of a tRNA (representing the mentioned minihelix) is discussed to be a relic of the RNA world, when genomes were build up of RNA or RNA-like polymers [87, 88]. A stem-loop structure at the 3′-end of these transcripts might have been the substrate for the addition of a CCA marker tag that introduced a replication start signal for these molecules [89]. This genomic tag was then used as a starting template for RNA replication, guiding the replicase to initiate polymerization with one or two G residues that base-pair with the CCA sequence [88]. Such a GG replication start has the advantage that the G:C base pairs have more hydrogen bonds than A:U interactions, stabilizing thereby the primer/template interaction. When later on modern telomerases evolved, these enzymes possibly "borrowed" the CCA tag for the endogenous RNA template that is still rich in C and A residues [89]. The original participants,
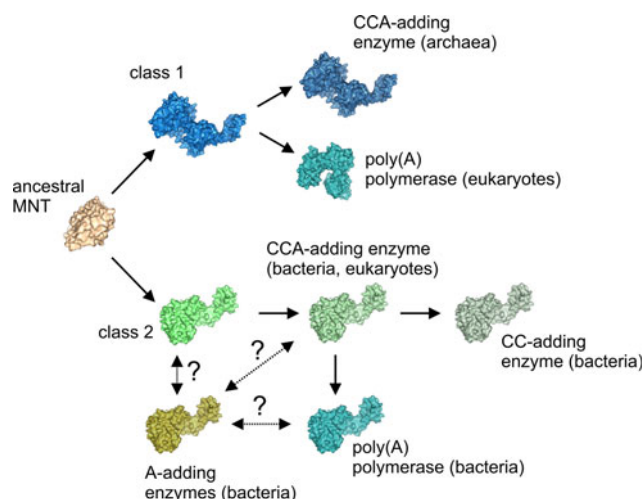

Fig. 8 Evolution of class I and class II CCA-adding enzymes as well as poly(A) polymerases. Starting from a primitive minimal nucleotidyltransferase MNT, class I and class II enzymes emerged rather early by the addition of functional domains. Class I split into two subgroups consisting of archaeal CCA-adding enzymes and eukaryotic poly(A) polymerases. In class II, however, CCA-adding enzymes probably emerged first and evolved then into enzymes with different (poly(A) polymerases) or partial activities (CC-adding enzymes). The evolution of A-adding enzymes is not understood (indicated by the *dashed arrows* and the *question marks*). It is possible that these enzymes present the ancestral state, leading to CCA-adding enzymes and poly(A) polymerases. Vice versa, it is also possible that one of these activities is the progenitor of the A-adding enzymes. Modified after [5]. Graphical presentations are based on the available pdb entries of the individual enzymes

however, the 3′-terminal RNA minihelix and the CCA-tagging activity, evolved into their modern functions in the protein world, the tRNA molecules and the two types of CCA-adding enzymes. Although these enzymes have obviously lost their original first jobs in the RNA world and had to be re-trained in their new employment, they still remain highly interesting and fascinating research objects that we still do not fully understand.

References

1. Aravind L, Koonin EV (1999) DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res 27:1609–1618
2. Holm L, Sander C (1995) DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. Trends Biochem Sci 20:345–347
3. Kwak JE, Wickens M (2007) A family of poly(U) polymerases. RNA 13:860–867
4. Ito J, Braithwaite DK (1991) Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res 19:4045–4057
5. Yue D, Maizels N, Weiner AM (1996) CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding

enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. RNA 2:895–908

6. Martin G, Keller W (2007) RNA-specific ribonucleotidyl transferases. RNA 13:1834–1849

7. Cabaniols JP, Fazilleau N, Casrouge A, Kourilsky P, Kanellopoulos JM (2001) Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase. J Exp Med 194:1385–1390

8. Li Z, Pandit S, Deutscher MP (1998) Polyadenylation of stable RNA precursors in vivo. Proc Natl Acad Sci USA 95:12158–12162

9. Aphasizhev R (2005) RNA uridylyltransferases. Cell Mol Life Sci 62:2194–2203

10. Trippe R, Guschina E, Hossbach M, Urlaub H, Luhrmann R, Benecke BJ (2006) Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA 12:1494–1504

11. Hartmann RK, Gossringer M, Späth B, Fischer S, Marchfelder A (2009) The making of tRNAs and more—RNase P and tRNase Z. Prog Mol Biol Transl Sci 85:319–368

12. Mörl M, Marchfelder A (2001) The final cut. The importance of tRNA 3′-processing. EMBO Rep 2:17–20

13. Schürer H, Schiffer S, Marchfelder A, Mörl M (2001) This is the end: processing, editing and repair at the tRNA 3′-terminus. Biol Chem 382:1147–1156

14. Wegrzyn G, Wegrzyn A (2008) Is tRNA only a translation factor or also a regulator of other processes? J Appl Genet 49:115–122

15. Sprinzl M, Cramer F (1979) The -C-C-A end of tRNA and its role in protein biosynthesis. Prog Nucleic Acid Res Mol Biol 22:1–69

16. Green R, Noller HF (1997) Ribosomes and translation. Annu Rev Biochem 66:679–716

17. Simonovic M, Steitz TA (2008) Peptidyl-CCA deacylation on the ribosome promoted by induced fit and the O3′-hydroxyl group of A76 of the unacylated A-site tRNA. RNA 14:2372–2378

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

19. Deutscher MP, Lin JJ, Evans JA (1977) Transfer RNA metabolism in *Escherichia coli* cells deficient in tRNA nucleotidyltransferase. J Mol Biol 117:1081–1094

20. Lizano E, Scheibe M, Rammelt C, Betat H, Mörl M (2008) A comparative analysis of CCA-adding enzymes from human and *E. coli*: differences in CCA addition and tRNA 3′-end repair. Biochimie 90:762–772

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

22. Clark JM (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res 16:9677–9686

23. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res 15:8783–8798

24. Deutscher MP (1982) tRNA nucleotidyltransferase. In: Boyer PD (ed) The enzymes. Academic Press, New York, pp 183–215

25. Masiakowski P, Deutscher MP (1980) Dissection of the active site of rabbit liver tRNA nucleotidyltransferase. Specificity and properties of subsites for donor nucleotide triphosphates. J Biol Chem 255:11240–11246

26. Masiakowski P, Deutscher MP (1980) Dissection of the active site of rabbit liver tRNA nucleotidyltransferase. Specificity and properties of the tRNA and acceptor subsites determined with model acceptor substrates. J Biol Chem 255:11233–11239

27. Vörtler S, Mörl M (2010) tRNA-nucleotidyltransferases: highly unusual RNA polymerases with vital functions. FEBS Lett 584(2):297–302

28. Xiong Y, Li F, Wang J, Weiner AM, Steitz TA (2003) Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. Mol Cell Biochem 12:1165–1172

29. 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

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

31. Cho HD, Sood VD, Baker D, Weiner AM (2008) On the role of a conserved, potentially helix-breaking residue in the tRNA-binding alpha-helix of archaeal CCA-adding enzymes. RNA 14(7):1284–1289

32. 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

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

34. 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

35. Cho HD, Chen Y, Varani G, Weiner AM (2006) A model for C74 addition by CCA-adding enzymes: C74 addition, like C75 and A76 addition, does not involve tRNA translocation. J Biol Chem 281:9801–9811

36. 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

37. Toh Y, Numata T, Watanabe K, Takeshita D, Nureki O, Tomita K (2008) Molecular basis for maintenance of fidelity during the CCA-adding reaction by a CCA-adding enzyme. EMBO J 27:1944–1952

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

39. Yue D, Weiner AM, Maizels N (1998) The CCA-adding enzyme has a single active site. J Biol Chem 273:29693–29700

40. Steitz TA (1998) A mechanism for all polymerases. Nature 391:231–232

41. Steitz TA, Smerdon SJ, Jager J, Joyce CM (1994) A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science 266:2022–2025

42. Martin G, Doublie S, Keller W (2007) Determinants of substrate specificity in RNA-dependent nucleotidyl transferases. Biochim Biophys Acta 1779(4):206–216

43. Schimmel P, Yang XL (2004) Two classes give lessons about CCA. Nat Struct Mol Biol 11:807–808

44. Augustin MA, Reichert AS, Betat H, Huber R, Mörl M, Steegborn C (2003) Crystal structure of the human CCA-adding enzyme: insights into template-independent polymerization. J Mol Biol 328:985–994

45. 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

46. 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(21):3353–3365

47. Martin G, Keller W (2004) Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. RNA 10:899–906

48. Li Z, Sun Y, Thurlow DL (1997) RNA minihelices as model substrates for ATP/CTP:tRNA nucleotidyltransferase. Biochem J 327:847–851

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

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

51. Cho HD, Verlinde CL, Weiner AM (2007) Reengineering CCA-adding enzymes to function as (U, G)- or dCdCdA-adding enzymes or poly(C, A) and poly(U, G) polymerases. Proc Natl Acad Sci USA 104:54–59

52. 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

53. Kim S, Liu C, Halkidis K, Gamper HB, Hou YM (2009) Distinct kinetic determinants for the stepwise CCA addition to tRNA. RNA 15:1827–1836

54. 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

55. McGann RG, Deutscher MP (1980) Purification and character-ization of a mutant tRNA nucleotidyltransferase. Eur J Biochem 106:321–328

56. Zhu LQ, Cudny H, Deutscher MP (1986) A mutation in Esche-richia coli tRNA nucleotidyltransferase that affects only AMP incorporation is in a sequence often associated with nucleotide-binding proteins. J Biol Chem 261:14875–14877

57. Hegg LA, Kou M, Thurlow DL (1990) Recognition of the tRNA-like structure in tobacco mosaic viral RNA by ATP/CTP:tRNA nucleotidyltransferases from Escherichia coli and Saccharomyces cerevisiae. J Biol Chem 265:17441–17445

58. Cho HD, Tomita K, Suzuki T, Weiner AM (2002) U2 small nuclear RNA is a substrate for the CCA-adding enzyme (tRNA nucleotidyltransferase). J Biol Chem 277:3447–3455

59. Williams MA, Johzuka Y, Mulligan RM (2000) Addition of non-genomically encoded nucleotides to the 3′-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently ter-minate with CCA. Nucleic Acids Res 28:4444–4451

60. Jin Y, Bian T (2004) Nontemplated nucleotide addition prior to polyadenylation: a comparison of Arabidopsis cDNA and geno-mic sequences. RNA 10:1695–1697

61. Zandueta-Criado A, Bock R (2004) Surprising features of plastid ndhD transcripts: addition of non-encoded nucleotides and polysome association of mRNAs with an unedited start codon. Nucleic Acids Res 32:542–550

62. Franze de Fernandez MT, Hayward WS, August JT (1972) Bacterial proteins required for replication of phage Q ribonucleic acid. Purification and properties of host factor I, a ribonucleic acid-binding protein. J Biol Chem 247:824–831

63. Scheibe M, Bonin S, Hajnsdorf E, Betat H, Mörl M (2007) Hfq stimulates the activity of the CCA-adding enzyme. BMC Mol Biol 8:92

64. Lee T, Feig AL (2008) The RNA binding protein Hfq interacts specifically with tRNAs. RNA 14:514–523

65. Kufel J, Allmang C, Verdone L, Beggs JD, Tollervey D (2002) Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. Mol Cell Biol 22:5248–5256

66. Hajnsdorf E, Regnier P (2000) Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. Proc Natl Acad Sci USA 97:1501–1505

67. Mohanty BK, Maples VF, Kushner SR (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. Mol Microbiol 54:905–920

68. Aravind L, Koonin EV (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci 23:469–472

69. Yakunin AF, Proudfoot M, Kuznetsova E, Savchenko A, Brown G, Arrowsmith CH, Edwards AM (2004) The HD domain of the Escherichia coli tRNA nucleotidyltransferase has 2′,3′-cyclic phosphodiesterase, 2′-nucleotidase, and phosphatase activities. J Biol Chem 279:36819–36827

70. Soukup GA, Breaker RR (1999) Relationship between internu-cleotide linkage geometry and the stability of RNA. RNA 5:1308–1325

71. Thompson JE, Venegas FD, Raines RT (1994) Energetics of catalysis by ribonucleases: fate of the 2′,3′-cyclic phosphodiester intermediate. Biochemistry 33:7408–7414

72. Bralley P, Chang SA, Jones GH (2005) A phylogeny of bacterial RNA nucleotidyltransferases: Bacillus halodurans contains two tRNA nucleotidyltransferases. J Bacteriol 187:5927–5936

73. Bralley P, Cozad M, Jones GH (2009) Geobacter sulfurreducens contains separate C- and A-adding tRNA nucleotidyltransferases and a poly(A) polymerase. J Bacteriol 191:109–114

74. Tomita K, Weiner AM (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

75. Tomita K, Weiner AM (2002) Closely related CC- and A-adding enzymes collaborate to construct and repair the 3′-terminal CCA of tRNA in Synechocystis sp. and Deinococcus radiodurans. J Biol Chem 277:48192–48198

76. Lizano E, Schuster J, Müller M, Kelso J, Mörl M (2007) A splice variant of the human CCA-adding enzyme with modified activity. J Mol Biol 366:1258–1265

77. Carpousis AJ (2007) The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. Annu Rev Microbiol 61:71–87

78. O'Hara EB, Chekanova JA, Ingle CA, Kushner ZR, Peters E, Kushner SR (1995) Polyadenylylation helps regulate mRNA decay in Escherichia coli. Proc Natl Acad Sci USA 92:1807–1811

79. Symmons MF, Williams MG, Luisi BF, Jones GH, Carpousis AJ (2002) Running rings around RNA: a superfamily of phosphate-dependent RNases. Trends Biochem Sci 27:11–18

80. Raynal LC, Krisch HM, Carpousis AJ (1998) The Bacillus subtilis nucleotidyltransferase is a tRNA CCA-adding enzyme. J Bacteriol 180:6276–6282

81. Sohlberg B, Huang J, Cohen SN (2003) The Streptomyces coe-licolor polynucleotide phosphorylase homologue, and not the putative poly(A) polymerase, can polyadenylate RNA. J Bacteriol 185:7273–7278

82. Lisitsky I, Klaff P, Schuster G (1996) Addition of destabilizing poly (A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA. Proc Natl Acad Sci USA 93:13398–13403

83. Yehudai-Resheff S, Schuster G (2000) Characterization of the E. coli poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence. Nucleic Acids Res 28:1139–1144

84. Patthy L (1991) Modular exchange principles in proteins. Curr Opin Struct Biol 1:351–361

85. Patthy L (1996) Exon shuffling and other ways of module exchange. Matrix Biol 15:301–310 discussion 311–312

86. Riley M, Labedan B (1997) Protein evolution viewed through Escherichia coli protein sequences: introducing the notion of a structural segment of homology, the module. J Mol Biol 268:857–868

87. 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

88. Maizels N, Weiner AM (1999) The genomic tag hypothesis: what molecular fossils tell us about the evolution of tRNA. In: Gesteland RF, Cech TR, Atkins JF (eds) The RNA world. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, pp 79–111

89. 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 196:327–328 discussion 329–330

90. Bard J, Zhelkovsky AM, Helmling S, Earnest TN, Moore CL, Bohm A (2000) Structure of yeast poly(A) polymerase alone and in complex with 3′-dATP. Science 289:1346–1349