

# Unusual evolution of a catalytic core element in CCA-adding enzymes

Andrea Hoffmeier, Heike Betat, Alexander Bluschke, Robert Günther, Sandy Junghanns, Hans-Jörg Hofmann and Mario Mörl\*

Institute for Biochemistry, University of Leipzig, Brüderstr. 34, 04103 Leipzig, Germany

Received September 14, 2009; Revised March 1, 2010; Accepted March 3, 2010

## ABSTRACT

CCA-adding enzymes are polymerases existing in two distinct enzyme classes that both synthesize the C-C-A triplet at tRNA 3'-ends. Class II enzymes (found in bacteria and eukaryotes) carry a flexible loop in their catalytic core required for switching the specificity of the nucleotide binding pocket from CTP- to ATP-recognition. Despite this important function, the loop sequence varies strongly between individual class II CCA-adding enzymes. To investigate whether this loop operates as a discrete functional entity or whether it depends on the sequence context of the enzyme, we introduced reciprocal loop replacements in several enzymes. Surprisingly, many of these replacements are incompatible with enzymatic activity and inhibit ATP-incorporation. A phylogenetic analysis revealed the existence of conserved loop families. Loop replacements within families did not interfere with enzymatic activity, indicating that the loop function depends on a sequence context specific for individual enzyme families. Accordingly, modeling experiments suggest specific interactions of loop positions with important elements of the protein, forming a lever-like structure. Hence, although being part of the enzyme's catalytic core, the loop region follows an extraordinary evolutionary path, independent of other highly conserved catalytic core elements, but depending on specific sequence features in the context of the individual enzymes.

## INTRODUCTION

All tRNA molecules carry the invariant sequence CCA at their 3'-end, representing an essential element for

aminoacylation and interaction of the tRNA with the ribosome during protein synthesis (1). The enzymes responsible for the generation and maintenance of this sequence are the ATP(CTP):tRNA nucleotidyltransferases or CCA-adding enzymes—unique polymerases that synthesize this sequence at high fidelity without requiring a nucleic acid as a template. Instead, amino acids in the NTP binding pocket form hydrogen bonds with the Watson–Crick edge of the incoming nucleotides. Together with the tRNA primer, this interaction defines the nucleotide specificity of the active site (2–5). Based on the conserved sequence patterns in the catalytic core, CCA-adding enzymes are divided into two distinct classes, where the archaeal enzymes are assigned to class I and eubacterial and eukaryotic enzymes to class II (6). These groups share only little sequence homology, but their overall architecture is defined by four major domains (head, neck, body and tail) of similar dimensions (7–10).

For class II CCA-adding enzymes, crystal structures of prokaryotic (*Bacillus stearothermophilus*, *Thermotoga maritima*) and eukaryotic members (*Homo sapiens*) are available (7,8,11). In the catalytic core located in the head domain, these enzymes carry a highly flexible region of 10–12 amino acids. Interestingly, this region is an essential enzyme element for the addition of the terminal A residue to the tRNA primer, as point mutations and deletions within this region dramatically interfere with A-incorporation (12–14). The addition of the two C residues, however, remained unaffected (13,15). The same loop element plays an important role in the evolution and functionality of a subset of closely related class II nucleotidyltransferases with partial activities. In several bacterial species, one enzyme exclusively adds two C residues to the tRNA 3'-end (CC-adding enzyme), and a second enzyme (A-adding enzyme) incorporates the terminal A, completing thereby the sequence (13,16–18). Interestingly, A-adding enzymes carry a similar flexible loop element that is required for the A-incorporation,

\*To whom correspondence should be addressed. Tel: +49 0341 9736911; Fax: +49 0341 9736919; Email: moerl@uni-leipzig.de

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

and amino acid replacements in this loop lead to a strong reduction in the enzymes' activity (19). In the CC-adding enzymes, on the other hand, the loop region was deleted during evolution, reducing the enzymatic activity to the addition of two C residues. This scenario is corroborated by the fact that the CC-adding enzyme of *Bacillus halodurans* regains the A-adding activity when it carries the inserted flexible loop of a closely related CCA-adding enzyme (13).

Due to such an important function for CCA-addition, one would expect that this loop element is highly conserved at the sequence level, similar to the other motif elements of the catalytic core of class II CCA-adding enzymes. However, sequence alignments indicate that this is not the case, and the loop sequences of several bacterial and eukaryotic class II CCA-adding enzymes vary dramatically, showing no co-evolution with other elements of the active site. This low conservation as well as the flexibility of this element indicates a possible function as a passive hinge allowing individual movements of N- and C-terminal enzyme regions during catalysis. However, a loop replacement by a likewise flexible glycine spacer was not compatible with full enzymatic activity, indicating a more complex loop function (12). Here, this unusual feature was investigated by reciprocal loop exchanges between several class II CCA-adding enzymes. Interestingly, the loops of enzymes from *H. sapiens*, *Escherichia coli* and *B. stearothermophilus* were not compatible and led to reduced enzymatic activities in the chimeras. A detailed phylogenetic analysis of class II CCA-adding enzymes revealed the existence of individual loop motifs, which are conserved only within the different bacterial and eukaryotic branches of the phylogenetic tree. Correspondingly, loop replacements within the individual families resulted in fully active CCA-adding enzyme chimeras, supporting the functional and evolutionary relevance of these loop families.

Hence, the presented data demonstrate the extraordinary evolution of this region in class II CCA-adding enzymes. Although being an important element for A-addition, it evolves independently of the highly conserved catalytic core motifs. However, the observed sequence families indicate that the loop element has certain evolutionary restrictions, probably depending on interacting enzyme regions within or close to the catalytic core.

## MATERIALS AND METHODS

### Construction of recombinant clones

cDNAs of CCA-adding enzyme from *H. sapiens*, *E. coli* and *B. stearothermophilus* were cloned as described (20,21). Point mutations leading to single amino acid exchanges were introduced in the corresponding cDNAs using the QuickChange site-directed mutagenesis procedure (Stratagene). Complete loop replacements were constructed according to Just *et al.* (12). The coding regions for loop sequences of *Drosophila melanogaster* and *Wigglesworthia glossinidia* were synthesized as

overlapping oligonucleotides and amplified by overlap extension polymerase chain reaction (PCR).

### Recombinant protein expression and purification

Wild-type as well as loop chimeras of CCA-adding enzymes were expressed in *E. coli* BL21(DE3) (Novagen). Cells were grown at 37°C in 500 ml LB medium containing 30 µg/ml kanamycin. Expression was induced at an OD<sub>600</sub> of 0.6 by addition of IPTG to a final concentration of 1 mM. After 3–4 h incubation at 37°C, cells were harvested by centrifugation, lysed by lysozyme treatment and sonication in ice-cold 20 mM Tris/HCl, pH 7.6, 0.5 M NaCl, 10 mM MgCl<sub>2</sub>, 5 mM imidazole and 100 µg/ml lysozyme. After centrifugation for 30 min at 24 000g at 4°C, the soluble recombinant proteins (carrying an N-terminal His<sub>6</sub> tag) were purified by affinity chromatography on a 1-ml HiTrap Chelating Sepharose column (GE Healthcare) and eluted with 500 mM imidazole (Supplementary Data, Figure 7). Fractions containing the enzymes were identified by SDS-PAGE and dialyzed against 20 mM Tris/HCl, pH 7.6, 0.2 M NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT and 10 % glycerol. Purified proteins were stored in the presence of 40% (v/v) glycerol at –20°C until use.

### Preparation of RNA substrate

Yeast tRNA<sup>Phe</sup> lacking the CCA-terminus or ending with a partial CCA-end (tRNA<sup>Phe</sup>-CC) was prepared as previously described (22,23). Except for substrates required for kinetic analyses, α-<sup>32</sup>P-ATP was added to the transcription mix to obtain internally labeled transcripts. RNA molecules were purified by denaturing polyacrylamide gel electrophoresis. Bands were cut out with a sterile blade, and the RNA was eluted by incubation in water at 4°C overnight. Transcripts were ethanol-precipitated in the presence of glycogen and redissolved in water.

### Enzyme activity assays

Two to four picomoles of <sup>32</sup>P-labeled tRNA substrate were incubated with 2–50 ng of recombinant enzyme in the presence of all four NTPs (1 mM each) in 30 mM HEPES/KOH, pH 7.6, 6 mM MgCl<sub>2</sub>, 30 mM KCl, 2 mM DTT in a total volume of 20 µl for 5–30 min at 30°C. The reaction was stopped by ethanol precipitation. The products were separated by electrophoresis on a 10% polyacrylamide gel containing 8 M urea and visualized by autoradiography.

### Kinetic analyses

For steady-state kinetic assays with seven data points, enzyme variants were tested at 30°C in a 10-µl reaction volume with CCA-addition buffer, 3 µM yeast tRNA<sup>Phe</sup> (lacking the CCA-terminus or ending with CC) and 1–3 µCi α-<sup>32</sup>P-ATP or CTP (3000 Ci/mmol) included as a label. ATP or CTP was titrated between 0.005 and 0.6 mM. The reactions were stopped after 10 min by adding 20 µl EDTA (50 mM) and spotted to DE81 filter papers (Whatman). Filters were washed with 100 ml buffer containing 0.3 M NH<sub>4</sub>-formate and 10 mM

pyrophosphate on a filter manifold (Amersham) and measured in a scintillation counter. Kinetic parameters of three to four independent experiments were analyzed using *GraphPadPrism* (curve fitting by nonlinear regression). Due to the solution properties of RNA transcripts, the obtained  $k_{\text{cat}}$  and  $K_{\text{M}}$  values are apparent values because the tRNA was not used at excessive saturating amounts when NTPs were titrated according to the literature (19,20,24,25). Additional kinetic parameters are indicated in Table 1 (Supplementary Data).

### Sequence analysis of reaction products

The 3'-ends of the isolated reaction products were analyzed as described (20).

### Computational analysis

Protein alignments were performed using *ClustalW* (<http://www.ebi.ac.uk/Tools/clustalw/>) with default parameters (26). Graphical presentation of conserved positions in the alignment was done using *Jalview* (27). For protein secondary structure prediction, the programs *Psipred* and *PredictProtein* were used (28–30). Sequence alignments of human and *B. stearothermophilus* CCA-adding enzymes were further adjusted by hand according to crystal structure overlays (13). A comprehensive protein data base search for unidentified CCA-adding enzyme sequences was performed using the genomic *BLAST* database at *NCBI* ([http://www.ncbi.nlm.nih.gov/sutils/genom\\_table.cgi](http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi)) and the *ExpASy BLAST* service (<http://www.expasy.org/tools/blast>). Sequence logos of the flexible loop regions were generated using *WebLogo* (31).

### Molecular modeling

The modeling studies of the flexible loop were performed using the program MOE 2008.10 (Chemical Computing Group, Inc. Montreal, Canada). The crystal structure of the human class II CCA-adding enzyme [pdb-entry 1OU5 (7)] served as the starting structure. Using the corresponding pdb-file from the RCSB data base (32), the residues of the flexible loops (positions 99–110) were modeled employing the 'Homology Model' module of the program. Hydrogen atoms were added and charges were assigned based on the Amber99 force field as incorporated in MOE. Chain A of the crystal structure served as the template. For each of the investigated heterologous loop sequences, 10 independent models were constructed with the default parameters. The model with the best score was then selected for further studies.

The molecular simulation calculations were performed with the program Gromacs 3.3. (33), employing the OPLS-AA/L force field (34). The TIP4P water model was chosen to consider solvent effects. In all cases, the protein was embedded in a dodecahedral box with a minimal distance to the edge of 7 Å. To ensure a neutral net charge of the system, counter ions were added using the genion program, which is part of the Gromacs package. The electrostatic interactions were calculated using the particle-mesh Ewald summation (PME). Coulomb and Lennard-Jones interactions were truncated

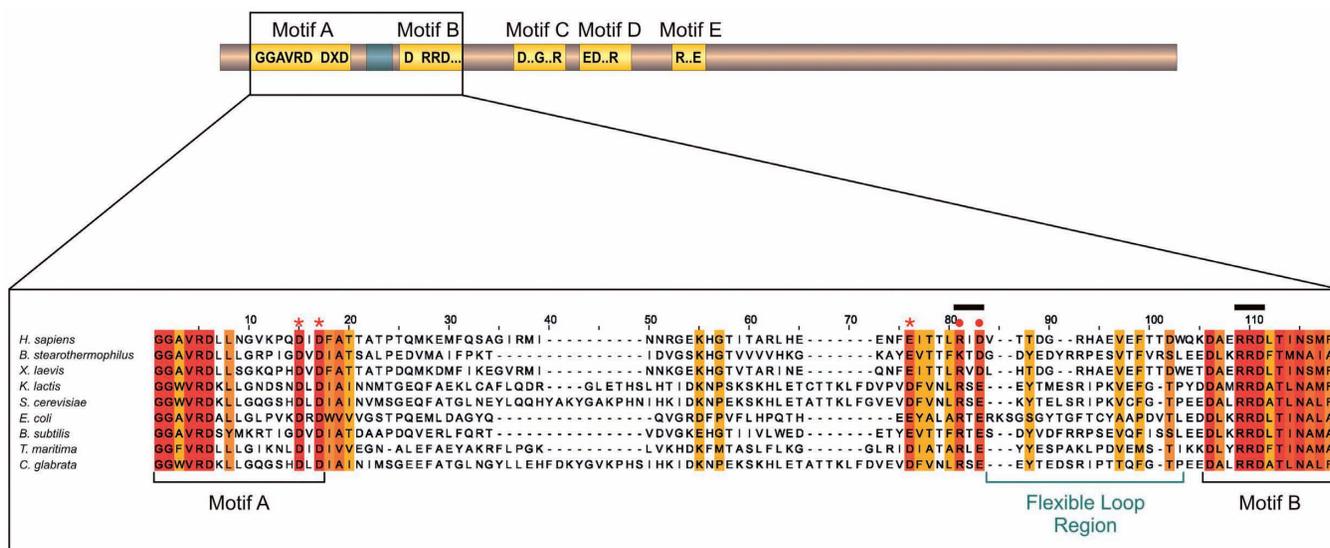
at 10 Å. Non-bonded pair interactions were updated every five steps. After a short relaxation of solvent molecules and ions, the system was minimized for 1000 steps with the steepest descent algorithm. The productive MD runs were performed at constant temperature (300 K) and constant pressure (1 bar) employing a Berendsen thermostat (35) with a coupling of 0.1 ps to the thermostat and 0.5 ps to the barostat. Bonds to hydrogen atoms were constraint with the SHAKE algorithm (36). The integration time step was set to 2 fs. Conformational snapshots of the system were saved every 1 ps.

## RESULTS

### Loop element replacements in human and bacterial class II CCA-adding enzymes interfere with A-addition

Class II nucleotidyltransferases share a set of five highly conserved motifs in the catalytic core, located in the N-terminal part (8). A sixth element important for catalysis was identified recently and is located as a flexible loop between motifs A and B (13). Using the alignment analysis software *Jalview*, the degree of conservation of the loop region as well as the flanking motifs in biochemically well characterized class II CCA-adding enzymes were determined (27; Figure 1). While both motifs A and B show the typical high level of conservation with several invariant positions, the sequence of the flexible loop element varies dramatically between individual enzymes. Yet, several point mutations as well as a replacement by an equally flexible (Gly)<sub>10</sub> spacer abolish the addition of the terminal A residue and restrict the enzymes' activity to CC-incorporation (12–14). Hence, these nonconserved loop elements must contain certain positions that are required for the full nucleotide incorporating activity.

One possibility is that these loops, although consisting of different individual sequences, fold up into a conserved and defined structure that is required for switching the enzymes specificity from C- to A-addition during polymerization. Hence, the loops would represent structural entities that should be compatible with each other and functional in a different enzymatic context. To test this idea, reciprocal exchanges of the region located between the conserved flanking motifs A and B described above were performed. As the CCA-adding enzymes of *H. sapiens*, *E. coli* and *B. stearothermophilus* are studied in most detail, these proteins were selected as candidates for loop replacements (Figure 2A). In order to reduce the risk of incompatible loop transplantations resulting in misfolded and inactive chimeric enzymes, the fusion sites were not positioned at the immediate N- and C-terminal borders of the loop sequences, but were shifted to the nearby highly conserved motifs flanking the loop region. The upstream located BxA motif (Figure 1) is represented by an invariant basic residue B followed by any amino acid x and a conserved acidic residue A and is probably involved in positioning the tRNA 3'-end for A-addition (11,13,19). As a downstream border, motif B was selected, carrying the RRD signature as a suitable fusion position (Figure 1). Using this strategy, residues 100–117 were replaced in the human



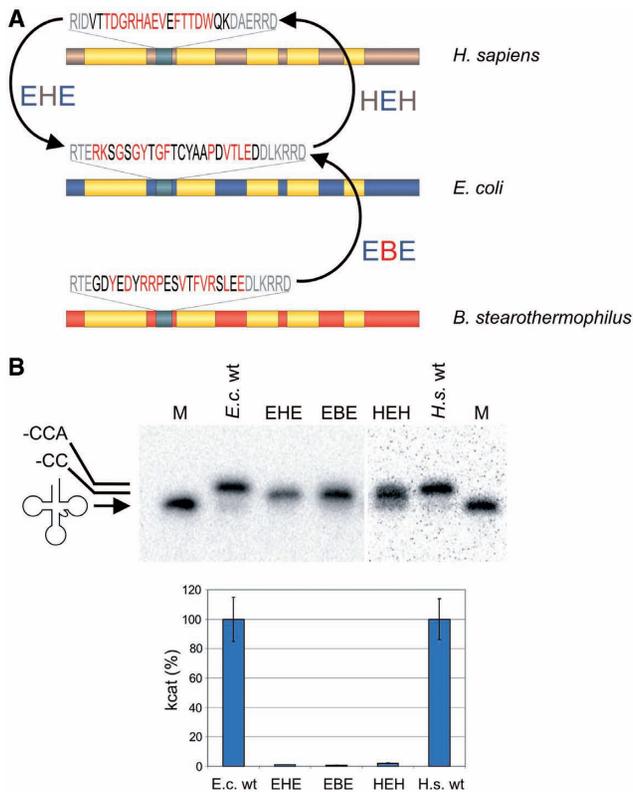
**Figure 1.** Conservation analysis of the region surrounding the flexible loop in biochemically well-characterized CCA-adding enzymes. Using the *Jalview* software, the conservation levels of individual amino acid positions are indicated by colors [from dark red (high) to light orange (low)]. The flexible loop region shows almost no conservation, while motifs A and B carry highly invariant positions. A similar level of conservation was found for the catalytically active carboxylates (labeled by red asterisks) and the basic/acidic motif BxA (indicated by red dots) that is probably involved in primer positioning. The black bars indicate the two highly conserved signatures (upstream of the loop: BxA; downstream: RRD in motif B) as selected fusion positions for the reciprocal exchanges of the loop regions.

enzyme by the corresponding part of the *E. coli* enzyme (positions 66–87), leading to the chimera HEH (H, human enzyme N-terminus; E, *E. coli* flexible loop; H, human enzyme C-terminus). Furthermore, the region in the *E. coli* enzyme was replaced by either the human loop element (representing the reciprocal experiment, chimera EHE) or by the *B. stearothermophilus* part (resulting in chimera EBE; here, the inserted *B. stearothermophilus* region contained also the BxA motif).

The resulting enzyme chimeras were recombinantly expressed in *E. coli*, purified by affinity chromatography and tested for CCA-adding activity on a radioactively labeled *in vitro* transcript of the yeast tRNA<sup>Phe</sup> (lacking the CCA-terminus) in the presence of NTPs. Subsequently, the reaction products were separated on a denaturing polyacrylamide gel and visualized by autoradiography (Figure 2B). Whereas the wild-type enzymes incorporated the complete CCA sequence as indicated by a pronounced mobility shift of the corresponding signal in the gel, the chimeric enzymes EHE, HEH and EBE showed a reduced activity and added only 2 nt to the tRNA substrate, leading to a different electrophoretic migration position. Sequence analysis of the reaction products revealed that these nucleotide additions correspond to an incorporation of only two C residues (data not shown), indicating that the incorporation of the third position (AMP) was affected. Hence, this A-addition was quantitatively characterized by the determination of the corresponding kinetic parameters (Figure 2B), using tRNA<sup>Phe</sup> ending with two C residues (tRNA<sup>Phe</sup>-CC) as a substrate. The obtained  $k_{\text{cat}}$  values are apparent values because it was not possible to use tRNA at saturating concentrations when NTPs were titrated. Nevertheless, the chosen conditions represent the standard procedure

to determine kinetic parameters of CCA-adding enzymes (19,20,24,25). The observed  $k_{\text{cat}}$  values of the wild-type enzymes (*H. sapiens* and *E. coli*) correspond to those described in the literature (20) and were set to 100% (Supplementary Data, Table 1). Compared to these values, the chimeras EHE, HEH, and EBE showed a 45- to 145-fold reduced  $k_{\text{cat}}$  for A-incorporation, ranging from  $0.69\% \pm 0.09$  (EBE) to  $1.03\% \pm 0.03$  (EHE) and  $2.18\% \pm 0.13$  (HEH), consistent with the observation that only 2 nt were incorporated in the reaction shown in Figure 2B. The corresponding  $K_{\text{M}}$  values, on the other hand, were in a range between 2.41 (EBE) and 108.4  $\mu\text{M}$  ATP (EHE), consistent with the  $K_{\text{M}}$  values of the loop donor enzymes (Supplementary Data, Table 1). However, as the A-adding activity is almost completely lost in these chimeras, these  $K_{\text{M}}$  values should be considered with care. Since the reaction products in Figure 2B show a highly efficient addition of the first two C residues, this incorporation reaction of the chimeric enzymes is obviously not affected by the loop replacements. Nevertheless, chimera EHE was chosen as a representative example for a kinetic analysis of C-incorporation and showed only a moderate 2-fold reduction in  $k_{\text{cat}}$  compared to the *E. coli* wt enzyme (data not shown).

Although the loop element inserted into the various enzyme variants apparently determines the  $K_{\text{M}}$  values for ATP, most of the loop replacements are not compatible with A-addition. Apparently, the composition of the individual loop region does not allow a simple exchange by a loop sequence of a different type, indicating that this element is not an isolated functional entity but probably interacts with some elements within the context of the individual CCA-adding enzymes. Consequently,

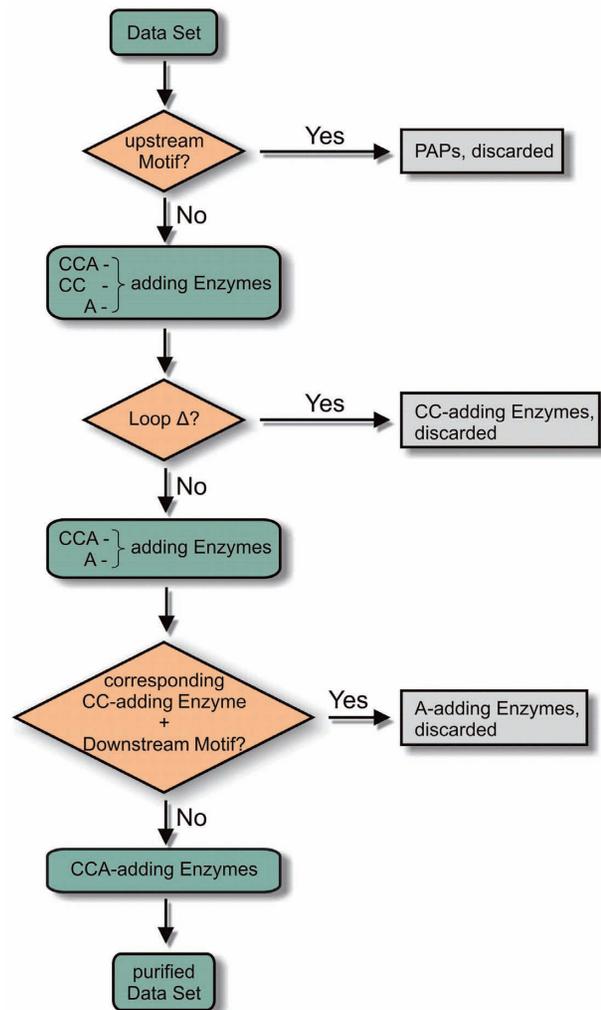


**Figure 2.** Loop exchanges between CCA-adding enzymes of *H. sapiens*, *E. coli* and *B. stearotherophilus*. (A) Arrows indicate the loop transplantations leading to chimeras HEH, EHE and EBE. Fusion positions were selected as described in the text. Characters labeled in red display conserved positions of each loop sequence. (B) (Upper panel) Nucleotide incorporation into a radioactively labeled tRNA substrate catalyzed by the chimeric enzymes. While wt enzymes of *E. coli* and *H. sapiens* showed a complete CCA-addition leading to a corresponding shift in the migration position of the products, the chimeras incorporated only two residues, but failed to add the terminal adenosine. (Lower panel) Relative  $k_{cat}$  values for ATP-incorporation of investigated enzymes. The values of the wt enzymes were set to 100%. The chimeras showed relative values close to 0%, demonstrating a dramatic reduction of the A-adding activity, while the CTP-incorporation was not affected (data not shown).

closely related class II CCA-adding enzymes should also carry related loop sequences. To identify such a possible conservation, a detailed and extensive sequence analysis of CCA-adding enzymes was performed.

### Conservation analysis of the flexible loop element in CCA-adding enzymes

Due to the numerous whole genome sequencing projects, an increasing number of sequences encoding putative CCA-adding enzymes is available. As the closely related bacterial poly(A) polymerases (PAPs) as well as CC- and A-adding enzymes share the same set of highly conserved motifs in the catalytic core (37), these enzymes are frequently incorrectly annotated as CCA-adding enzymes (38,39). However, recent analyses identified a subset of individual and highly specific sequence features that allow discrimination between these closely related enzyme forms (13,40). For a conservation analysis of the



**Figure 3.** Data set cleaning procedure for the identification of CCA-adding enzyme sequences. The starting data set of class II nucleotidyltransferases was retrieved by a genomic *BLAST* analysis of the *NCBI* database, using sequences of experimentally verified CCA-adding enzymes. According to the identity motifs of the individual types of class II nucleotidyltransferases, this original data set was purified from other entries (CC- and A-adding enzymes, poly(A) polymerases) and finally contained 339 sequences that carried all identity elements of genuine CCA-adding enzymes. This purified data set was used for further analysis.

flexible loop element identified in class II CCA-adding enzymes, a *BLAST* analysis was performed to retrieve a collection of enzyme sequences from the *NCBI* database ([http://www.ncbi.nlm.nih.gov/sutils/genom\\_table.cgi](http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi)), using sequences of experimentally verified CCA-adding enzymes as a query. In order to analyze exclusively sequences of genuine CCA-adding enzymes, the data set was purified from related poly(A) polymerases, CC- and A-adding enzymes according to the procedure described below (Figure 3). Bacterial poly(A) polymerases could be identified by the presence of an upstream motif [LIV]-[LIV]-G-[RK]-[RK]-F-x-[LIV]-h-[HQL]-[LIV], where x represents any residue and h a hydrophobic residue. This motif is located between motif A—the common signature of all nucleotidyltransferases—and the position of

the flexible loop. It is not found in CCA-, CC- or A-adding enzymes (40). The CC-adding enzymes were recognized by the existence of a deletion of the flexible loop element, which was described as an identity element for this type of enzymes (13). Finally, A-adding enzymes were also excluded from the data set. Here, two features were used to identify these nucleotidyltransferases. If an organism carried already a CC-adding enzyme (identified by the loop deletion), the second nucleotidyltransferase sequence is highly likely to represent an A-adding enzyme. Such sequences could be further identified by the existence of a downstream consensus motif sRxxxExxxhh (s, small residue; h, hydrophobic residue; x, any residue) which is located downstream of the catalytic core. This sequence is specific for PAPs and A-adding enzymes, while CC- and CCA-adding enzymes carry the consensus ERxxxExxxhh, where the N-terminal small residue (s) is replaced by glutamic acid (E) (40). Following this procedure, a total set of 339 individual sequences of class II CCA-adding enzymes was retrieved.

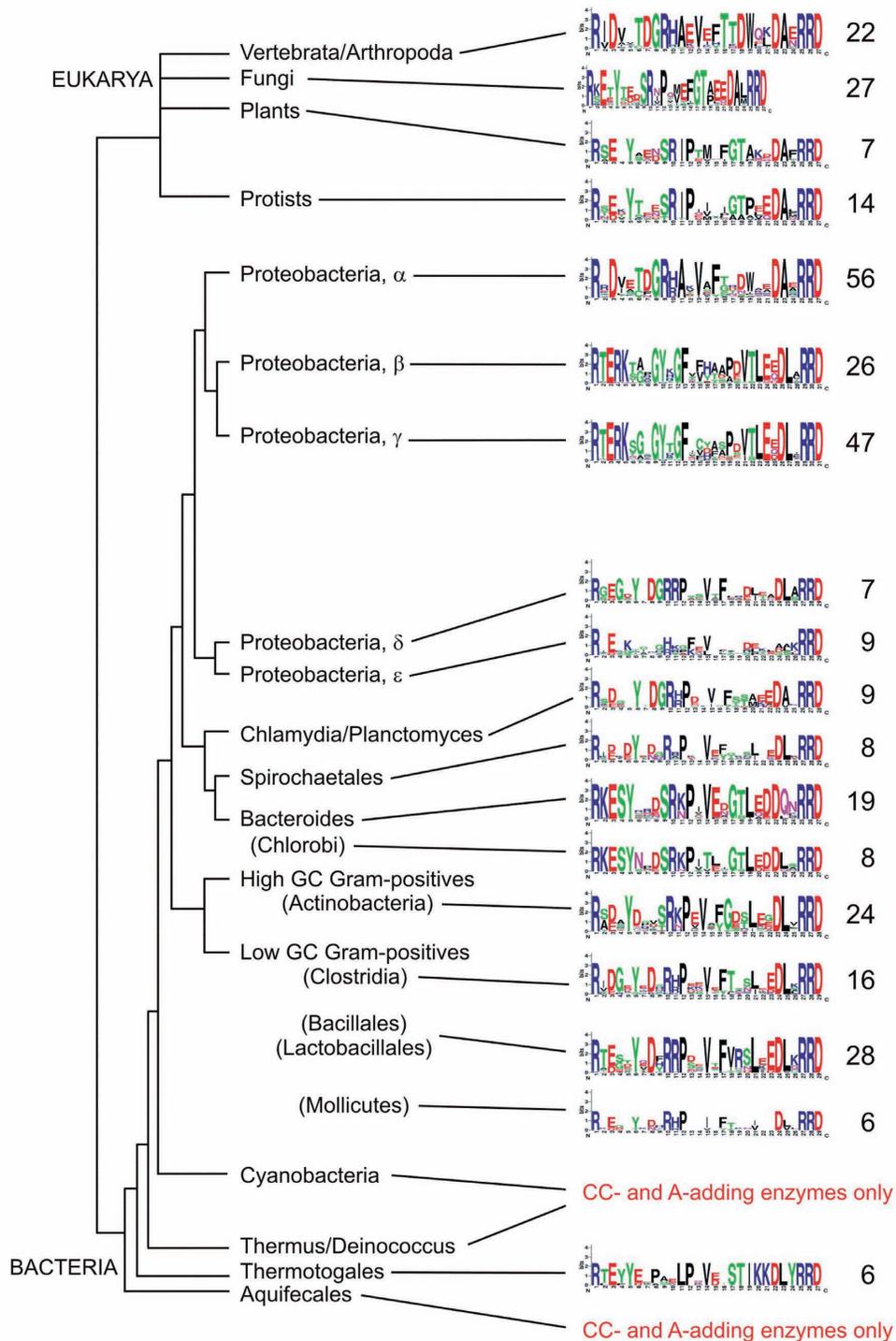
In a *ClustalW*-based multiple alignment of this data set, the analysis was restricted to the positions between two highly conserved elements flanking the loop structure. The N-terminally located border is represented by the BxA motif, while the conserved sequence RRD in motif B, involved in ribose recognition (8), was chosen as the C-terminal border (Figure 1). To investigate the extent of conservation of amino acid positions within the flexible loop, the resulting alignments of 18 genera were analyzed using the software *Weblogo* (<http://weblogo.berkeley.edu/>), which produces a graphical representation that indicates the level of conservation at each position by the size of the individual characters (31). This analysis revealed that the loop elements of each individual genus carry several moderately to highly conserved positions. In order to shed light on the phylogenetic relation, these sequence families were integrated into an adapted phylogenetic tree based on rRNA sequences (Figure 4; 40–42). Interestingly, phylogenetically closely related genera carry similar loop sequences. Furthermore, although certain positions show a rather high conservation across the analyzed genera, there is no general invariant amino acid residue within this region, indicating different modes of evolution for individual loop families.

#### Loop elements within a phylogenetic sequence family can be exchanged without affecting the overall enzymatic activity

In the experiments shown in Figure 2, loop elements of enzymes from *γ-Proteobacteria*, *Bacillales* and *Vertebrata/Arthropoda* were reciprocally exchanged, representing members of different and only distantly related loop sequence families (Figure 4). To analyze the loop sequence compatibility within a phylogenetic group, corresponding exchanges of individual loops were performed according to the strategy described above, starting immediately downstream of the BxA motif and ending upstream of the first highly conserved D residue of motif B. For the group of *γ-Proteobacteria*, the enzymes of *E. coli* and *W. glossinidia* were chosen. *W. glossinidia* is

an endosymbiotic Gram-negative bacterium isolated in the Tsetse fly gut, where the genome became available recently (43,44). In its CCA-adding enzyme, the region containing the flexible loop consists of 22 residues, where 12 positions differ from that of the *E. coli* enzyme, while only 10 positions are identical (positions 65–86 in *W. glossinidia* and 66–87 in *E. coli*; Figure 5A, left). Yet, when tested *in vitro* on a radioactively labeled tRNA<sup>Phe</sup> substrate, the resulting recombinant chimeric enzyme EWE (E, *E. coli* enzyme N-terminus, W, *W. glossinidia* loop region, E, *E. coli* enzyme C-terminus) showed a full CCA-adding activity, comparable to that of the wt *E. coli* enzyme, as both reaction products showed an identical migration position in the denaturing gel (Figure 5B, left). Furthermore, a second enzyme chimera was generated within the group of *Vertebrata/Arthropoda*, consisting of the loop element from *Drosophila melanogaster* (positions 159–176) inserted into the context of the CCA-adding enzyme of *H. sapiens*, replacing the positions 100–117 (HDH, nomenclature according to the ones above; Figure 5A, right). As in the case of the EWE chimera, this enzyme added a complete CCA triplet to the 3'-end of tRNA<sup>Phe</sup>, indicated by a migration position in the gel identical to that of the human wt enzyme (Figure 5B, right). Correspondingly, the kinetic analysis of both EWE and HDH chimeras revealed an only slightly reduced  $k_{cat}$  for A-incorporation compared to the wt enzymes (EWE: 1.8-fold reduction; HDH: 1.3-fold reduction; Figure 5C, Supplementary Data, Table 1). Furthermore, the  $K_M$  values of these chimeras were moderately increased compared to the human (3.8x) and the *E. coli* wt enzymes (3x), respectively. However, as the CCA-adding enzymes from *Drosophila* and *Wigglesworthia* are not available, it is not possible to compare the  $K_M$  values of the chimeras to those of the loop donor enzymes. Taken together, the observed CCA-adding activity of the EWE and HDH chimeras indicates that loop sequences of CCA-adding enzymes within a phylogenetic group are highly compatible, and the enzymes tolerate sequence replacements without affecting the overall catalysis. Nevertheless, the 5-fold reduction in  $k_{cat}/K_M$  (Supplementary Data, Table 1) implies that the inserted loop structures are not evolutionary adapted or fine-tuned to the foreign enzyme context.

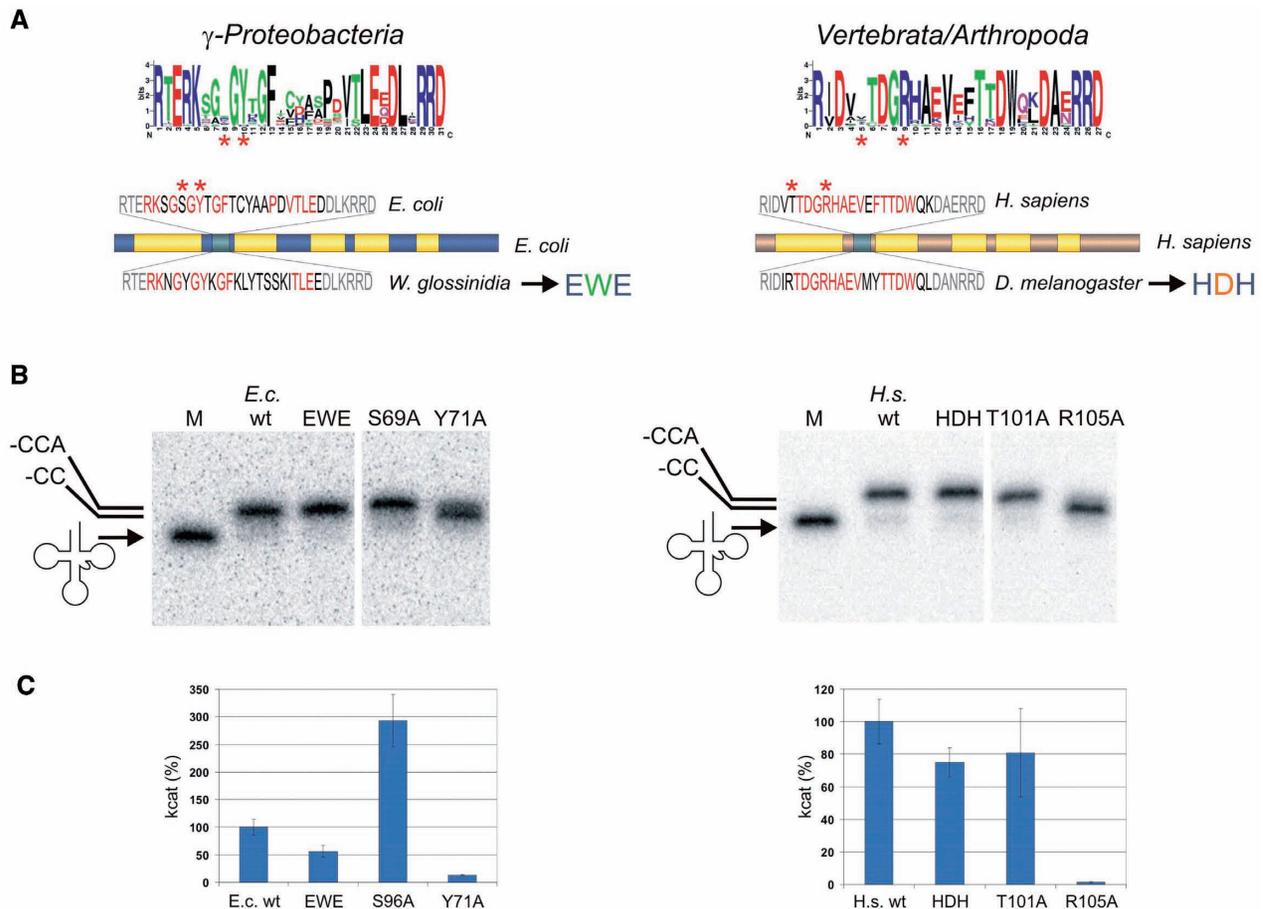
The involvement of the loop element in the addition of the terminal A residue was demonstrated by the introduction of several point mutations (G70D, G70A and Y71A) in the *E. coli* enzyme that almost completely abolished A-incorporation (12,14). The functional importance of these positions is in accordance with their high conservation in the loop sequence logo of *γ-Proteobacteria* (Figure 4). Furthermore, alanine replacements of amino acids D103, G104 and R105 within the loop region of the human enzyme led to a comparable effect and allow only the incorporation of two C residues (data not shown). Again, these residues represent conserved loop positions in the corresponding loop family (Figure 4). To address the question whether nonconserved positions in loop families are also important, position S69 in the *E. coli* enzyme and T101 in the human enzyme were replaced



**Figure 4.** Families of individual loop sequences. A phylogenetic tree based on rRNA sequences (41,42) was used to illustrate the loop consensus sequences identified for each of the presented genera. While some positions are found in several families, no amino acid residue is present at a certain position in all families, indicating that the loops do not contain absolutely conserved positions. The figures behind the sequence logos indicate the number of sequences of each genus that was used for this analysis. For the sequence alignment, *Clustal W* and *Weblogo* software was used. The analysis was restricted to the flexible loop region flanked by the conserved elements BxA (basic/acidic motif, N-terminal of the loop) and DxxRRD (motif B, C-terminal of the loop).

by alanine. The resultant recombinant enzymes were tested on the radioactively labeled tRNA<sup>Phc</sup> substrate and showed a complete CCA-addition, in contrast to the representative mutations Y71A (*E. coli*) and R105A

(human) at highly conserved positions that added only two C residues (Figure 5B). The analysis of kinetic parameters for the A-incorporation reaction revealed that the replacement of the conserved positions reduced the  $k_{cat}$



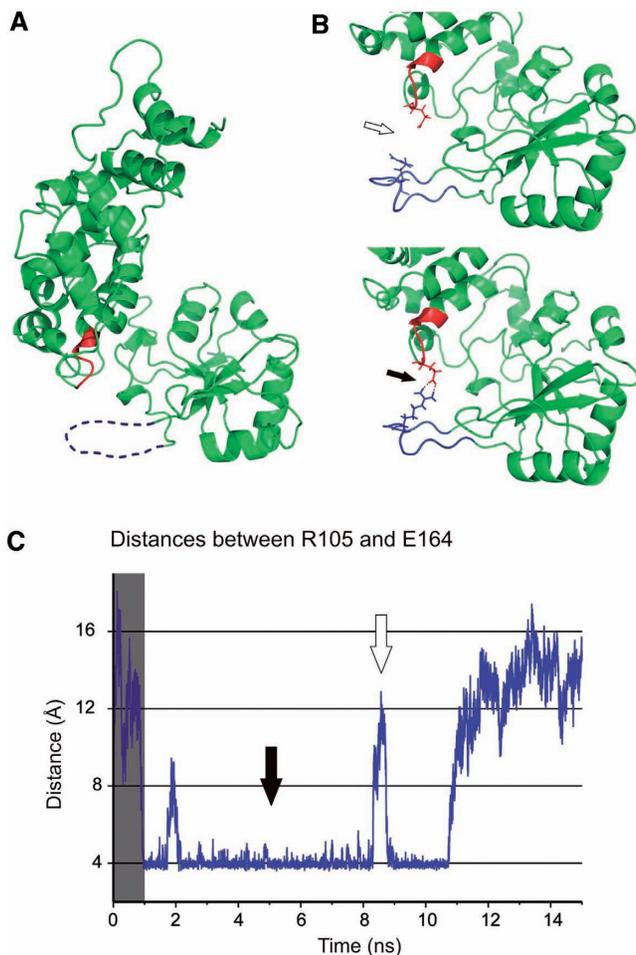
**Figure 5.** Loop replacements within sequence families. (A) Enzymes of *E. coli*/*W. glossinidia* and *H. sapiens*/*D. melanogaster* were chosen as representatives for  $\gamma$ -Proteobacteria and Vertebrata/Arthropoda, leading to loop chimeras EWE and HDH, respectively. Fusion positions were selected as in Figure 2. Conserved positions in the individual loop sequences are indicated in red, introduced point mutations S69A, Y71A (both in *E. coli*), T101A and R105A (both in *H. sapiens*) are labeled by asterisks. (B) Nucleotide addition activity of chimeras and enzymes carrying point mutations. Both EWE and HDH chimeras show a full CCA-adding activity, indicating compatible loop sequences of the parental wt enzymes that belong to the same sequence families. The point mutations at highly conserved positions (Y71A, R105A) dramatically interfere with ATP-addition, while CTP-incorporation remains unaffected. Correspondingly, alanine replacements at nonconserved positions do not influence the overall activity of the enzymes. (C) Comparative analysis of  $k_{cat}$  values of the mutant enzymes. While the point mutations demonstrate that conserved loop positions are essential for A-addition, the loop exchanges within sequence families affect the A-incorporation only slightly.

values to  $13.5\% \pm 0.9$  (7.5-fold) for Y71A in the *E. coli* enzyme and  $1.3\% \pm 0.5$  (77-fold) for R105A in the human enzyme, respectively (Figure 5C). The  $K_M$  values, however, changed only slightly. While the *E. coli* variant showed a  $K_M$  value almost identical to the wt ( $2.99 \mu\text{M}$  versus  $3.25 \mu\text{M}$ ), the  $K_M$  of the human version was even somewhat reduced ( $38.54 \mu\text{M}$  versus  $98.67 \mu\text{M}$ ). However, as these two protein forms are almost completely inactive in terms of A-addition (as some of the chimeras mentioned above), one should interpret these values very carefully. The mutations at the non-conserved positions, on the other hand, showed no significant decrease ( $80.7\% \pm 27.2$ ; human enzyme T101A) or even a threefold increase in A-adding activity ( $293.6\% \pm 47.6$ ; *E. coli* enzyme S69A). In agreement with this, the corresponding  $K_M$  values of the latter enzyme versions were only moderately affected (*E. coli* S69A:  $7.28 \mu\text{M}$  versus  $3.25 \mu\text{M}$  for the *E. coli* wt enzyme; *H. sapiens* T101A:  $78.5 \mu\text{M}$  versus  $98.67 \mu\text{M}$  for the human wt enzyme).

Taken together, although there are no general invariant amino acid positions common to all loop families, these data indicate that within the families, conserved positions are functionally important, while nonconserved positions can be exchanged without affecting the overall enzymatic catalysis.

#### Modeling studies on the flexible loop element

In order to find possible interaction partners for amino acid residues located within the loop, a computer-based modeling of a loop sequence inserted into the structure of a CCA-adding enzyme was performed. The crystal structure of the human enzyme (pdb-entry 1OU5) served as starting point (Figure 6A). The loop structure of this enzyme was modeled by a homology modeling approach as described in 'Materials and Methods' section. To investigate the dynamical behavior of this model, molecular dynamics simulations of the protein enclosed in a water box were performed. After a short initial phase of about



**Figure 6.** Molecular modeling of loop interactions. (A) Representation of the crystal structure of the human CCA-adding enzyme. The position of the flexible loop is indicated by the dashed blue line. The presentation is based on the pdb deposit number 1OU5 (7). (B) In the simulation, an arginine in the flexible loop (blue) forms a salt bridge to a glutamate residue of the nucleotide binding pocket (red), changing the enzyme from an open (upper panel) to a closed conformation (lower panel). (C) Time course of the simulation. After an initial phase of 1 ns (gray), the salt bridge with a distance of about 4 Å between R105 and E164 is formed, leading to a closed conformation of the enzyme (black filled arrow). Short interruptions of this interaction lead to the open conformation with a distance of 8–12 Å between these residues (open arrow), indicating the dynamic nature of this interaction.

1 ns, the conserved loop residue R105 formed a salt bridge to the first residue E164 of the amino acid template (EDxxR) in motif D (Figure 6B and C). Within the subsequent 15 ns of total simulation time, this interaction persisted for more than 8 ns, with only two short interruptions, indicating a rather stable and robust contact. In this closed conformation, the distance between the side chains of the arginine and the glutamate residues was 4 Å on average, representing a typical value for a salt bridge. However, the increase of the distance after ~11 ns demonstrates the dynamic nature of this loop element. The open conformation, in contrast, showed a distance of up to 16 Å between these residues.

In a second study, the modeled human loop was replaced by the *D. melanogaster* loop sequence, representing the

chimera HDH that showed an almost unaffected A-adding activity in the experiment (Figure 5). Similar to the original human enzyme, the heterologous loop formed a salt bridge between the corresponding arginine residue and the glutamate of the amino acid template in the molecular dynamics simulation (data not shown). Hence, both enzyme versions showing an A-adding activity form this interaction between loop and amino acid template in the modeling study. The functional importance of this interaction is also experimentally corroborated by the observation that the replacement of R105 by alanine in the human enzyme dramatically interfered with A-addition (Figure 5, Supplementary Table S1). In accordance with this finding, the replacement of the other interaction partner E164 by alanine also dramatically reduced  $k_{cat}$  according to the kinetic analysis, although this variant showed A-addition in previous endpoint determination studies (45; Hoffmeier, data not shown). However, such endpoint analyses do not allow any conclusions in terms of reaction efficiency. Furthermore, the  $K_M$  of this variant was 4.7-fold increased compared to the wt (see Supplementary Table S1), which might represent the involvement of this glutamate in nucleotide binding.

## DISCUSSION

The crystal structures of three members of class II tRNA nucleotidyltransferases are well described and display a very high structural similarity. Both CCA- and A-adding enzymes show a hook-like shape with a catalytic core located in the N-terminal part. This active site consists of five highly conserved elements (motifs A–E; Figure 1) with different functions, ranging from metal ion binding to nucleotide recognition and templating (7,8,19). Interestingly, between motifs A and B, these enzymes carry a short region containing 10–12 amino acids that seems to be highly disordered and flexible, as it is not resolved in most of the crystal structures (Figure 6A, indicated by the dashed line; 7,8,11,19). Furthermore, this element shows no conservation at the sequence level between several well-characterized CCA-adding enzymes (Figure 1). Yet, point mutations within this region dramatically reduce the enzyme's capacity to incorporate the terminal A residue of the CCA terminus in the tRNA substrate (12,14,19). In accordance with these observations, the CC-adding enzymes that lack an A-adding activity have a deletion of this region in common (13). Hence, the loop represents a functional element required for A-addition but shows remarkable sequence diversity and no well-defined tertiary structure. A possible role as an ATP binding pocket was discussed for the *E. coli* enzyme due to similarities to nucleotide-binding P-loop sequences (14), but the low sequence conservation excludes such a function (13).

Nevertheless, the mutational effects on the A-incorporation indicate that the individual loops carry positions that cannot be replaced randomly but must have an intrinsic function. Two scenarios for a possible mode of operation of this element are conceivable. The important residues of the loop could form restricted intra-loop

interactions with other positions within this isolated region. Hence, the loop would represent a discrete functional entity that does not interact with other elements of the enzyme. Consequently, exchanges of such loop entities between different enzyme families should be compatible with CCA-adding activity. However, the corresponding chimeras HEH, EHE and EBE are dramatically inhibited in A-addition, indicating that the context of the enzyme is important for a functional loop element. The alternative scenario is based on interactions of loop residues with positions located elsewhere in the enzyme body. Here, enzymes of closely related species with highly conserved sequences should also carry similar loop elements. The phylogenetic analysis of 339 individual CCA-adding enzymes supports this idea, showing individual loop sequence families according to the phylogenetic origin of the corresponding organisms (Figure 4). Accordingly, the loop sequences within families should form identical or very similar interactions to positions in the enzyme body, allowing reciprocal intra-family loop exchanges without dramatic loss in A-adding activity. This is supported by the second set of chimeric enzymes EWE and HDH, where loop regions were exchanged within the identified phylogenetic families. Both chimeras showed an almost unaffected A-incorporation, although the loop sequences differed at several positions (11 differences out of 22 between *E. coli* and *W. glossinidia* enzymes, six differences out of 18 between *H. sapiens* and *D. melanogaster* enzymes). Regarding the conserved positions in the consensus sequences of  $\gamma$ -Proteobacteria and Vertebrata/Arthropoda, however, only two residues varied in the two bacterial enzymes (the *W. glossinidia* enzyme carries S and I instead of the conserved P and V residues), while only one difference was detected between the human and the fly enzyme (the *D. melanogaster* enzyme carries a Y residue instead of the conserved F). Apparently, these positions are not functionally important, although they show a moderate conservation. The other existing invariant residues, however, can obviously find their corresponding interaction partners in the context of the foreign enzyme, thereby allowing a complete CCA-addition. The importance of some of these positions was further demonstrated by mutational replacement. Compared to alanine replacements of nonconserved residues (which did not affect A-incorporation or even enhanced this reaction), the mutations of conserved residues strongly reduced this reaction (Figure 5B and C). Furthermore, the  $K_M$  values of enzymes that still incorporate the terminal A residue (chimeras EWE, HDH as well as point mutations T101A in the human enzyme and S96A in the *E. coli* enzyme) are in the range of those of the wt enzymes, showing only a moderate (2- to 3-fold) increase (Supplementary Data, Table 1). When all chimeras (active as well as inactive in terms of A-addition) are compared to the parental wild-type proteins, a surprising consistency of the  $K_M$  values to the loop donor enzymes can be observed. The EHE variant, for example, has almost completely lost its A-incorporating activity. However, the  $K_M$  of this enzyme (108.4  $\mu$ M) is close to that of the human loop donor enzyme (98.67  $\mu$ M). Similarly, the EBE chimera

and the enzyme from *B. stearothermophilus* have a nearly identical  $K_M$  (2.41  $\mu$ M versus 3.99  $\mu$ M), which can be also observed for the HEH chimera (3.11  $\mu$ M) and the *E. coli* enzyme (3.25  $\mu$ M). Although these data seem to indicate a direct involvement of the loop in ATP binding, one should keep in mind that these enzymes do not show a detectable A-adding activity. Hence, the  $K_M$  values should be interpreted rather carefully. A further indication that these  $K_M$  values do not allow conclusions concerning the enzymatic activity comes from the fact that enzyme versions that differ dramatically in their A-adding activity can have an almost identical  $K_M$ . The human R105A variant has a  $K_M$  of 38.54  $\mu$ M, which is similar to that of the corresponding wt enzyme (98.67  $\mu$ M). Yet, the mutant enzyme has lost its A-adding activity. In line with this observation, the human E164A variant and the HDH chimera have similar  $K_M$  values (463  $\mu$ M and 373.6  $\mu$ M, respectively), but only the HDH enzyme shows an efficient A-incorporation (Supplementary Data, Table 1).

Nevertheless, an indirect effect on ATP binding is conceivable. The nucleotide binding pocket of the enzyme contains the amino acid template (motif D) for proper recognition of CTP and ATP. Amino acids within this motif form hydrogen bonds with the base moiety of the incoming nucleotides. A reorientation of these amino acids is required to switch from CTP- to ATP-binding, thereby enlarging the binding pocket so that it can accommodate the bigger adenosine base (8). This reorientation is obviously part of a conformational rearrangement of the whole enzyme during ATP-incorporation, as enzyme crystals readily dissolve upon soaking with ATP (19). Hence, in such an overall enzyme movement, an interaction with the loop element could form a lever or arm that is required for the observed reorientation of the templating amino acids in the nucleotide binding pocket. Indeed, the molecular dynamics simulation (Figure 6) indicates such an interaction of an arginine residue of the flexible loop and the glutamate of the EDxxR templating region in both fully active human wt enzyme and HDH chimera, forming a lever structure that pulls the amino acid template into the position required for recognition of ATP instead of CTP. An interruption of this R-E interaction would interfere with this reorientation of the binding pocket, so that ATP binding (and, consequently, incorporation) is inhibited. However, in such a scenario, the binding pocket would remain in a CTP-binding mode, and one would expect the addition of a third C residue instead of the terminal A. As this is not the case, it is more likely that—in addition to the lever function—the loop element directly contributes to catalysis of the A-addition. Recently, a possible interaction of the loop with the C75 primer end of the tRNA was described, which positions the 3'-OH group of the tRNA for the nucleophilic attack on the triphosphate moiety of the bound ATP (11). In the chimeras that do not incorporate ATP, this positioning is obviously not possible, as the inserted loop element is not compatible with the geometry of the catalytic core of the foreign enzyme.

A further indication for such incompatibilities is also observed in the intra-family loop chimeras that add a

complete CCA-terminus. The slight decrease in  $k_{\text{cat}}$  of the EWE and HDH enzymes indicates that these loops are not optimally adapted to the overall enzyme core, leading to a reduced efficiency as a lever required for the overall enzyme movement. Moreover, the observed R–E interaction cannot represent a general element found in all loop families, as the arginine residue is not present in all loop sequences (e.g.  $\beta$ - and  $\gamma$ -*Proteobacteria* enzymes do not possess this residue). Indeed, in the crystal structure of the *T. maritima* CCA-adding enzyme, a tyrosine residue in the loop was described that forms a hydrophobic interaction with an aspartic acid residue of the amino acid template that corresponds to the glutamate position in the R–E interaction found in our modeling experiment (11). Hence, since no absolutely invariant position common to all loop sequences could be identified, it seems that in certain enzymes other positions in the loop might be involved in similar interactions. Likewise, one cannot exclude that also residues outside of the amino acid template interact with the loop and form a lever structure. A prerequisite for such interactions is a co-evolution of the loop sequence with its binding partner(s) in the central body of the CCA-adding enzyme. This would explain the observed incompatibility of the loop insertion into an enzyme body of a different phylogenetic group. As the motifs of the catalytic core of an enzyme usually co-evolve in order to maintain and optimize the catalytic activity, it is a fascinating observation that the CCA-adding enzymes carry an element that—although involved in the proteins' functionality and reactivity—obviously does not follow this principle of co-evolution.

Furthermore, similar loop elements were identified in other polymerases as well. While these loops are also important for the catalytic activity of these enzymes, the actual mode of operation is still unclear. DNA polymerase  $\beta$ —an enzyme that is found in the same nucleotidyltransferase superfamily as the CCA-adding enzymes—also carries such a highly flexible region of similar length (46). Here, a deletion analysis revealed that this loop contributes to the fidelity of the polymerase involved in base excision repair. Within a specific template sequence context, the deletion variant of polymerase  $\beta$  produces frame shifts as well as A to C transitions, and the authors discuss a direct or indirect influence of the loop on the nucleotide binding pocket, stabilizing primer and template (46). A further member of the polymerase  $\beta$  superfamily that carries such a loop region with functional importance on catalysis is the human Pol  $\mu$ , an enzyme closely related to terminal deoxynucleotidyltransferase (TdT; 47). This enzyme has a catalytic core similar to that of polymerase  $\beta$  and contains a flexible element of 17 amino acids in length, located at the same distance from the catalytic DxD signature (motif A) as the corresponding element in the CCA-adding enzyme. Deleting this loop converted the enzyme from a template-independent polymerase into a DNA–template-dependent enzyme. Moreover, such flexible loop elements are not restricted to polymerases, but are also found in other enzymes, like the human hypoxanthine phosphoribosyltransferase (48), HIV-1 integrase (49), *E. coli* undecaprenyl-phosphate

synthase (50) and yeast pyruvate decarboxylase (51), where these regions are also involved in structural rearrangements.

The examples of polymerase  $\beta$  and Pol  $\mu$  indicate that the existence of flexible loop regions involved in polymerase templating and fidelity is not restricted to class II CCA-adding enzymes. Nevertheless, it is not known yet whether these various polymerase loop elements show a similar evolutionary pattern, independent of the catalytic core as found in the class II CCA-adding enzymes. It will be fascinating to see whether this is a specific feature of tRNA nucleotidyltransferases or a more general phenomenon that is also common to other nucleotide-incorporating enzymes.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank M. Grunow and A. Neuenfeldt for valuable discussions and S. Bonin and A. Krahl for expert technical assistance. Thanks to C. Teichmann for her help with the construction of some of the mutated enzyme clones.

## FUNDING

The Deutsche Forschungsgemeinschaft (Mo 634/2-4, Mo 634/3-3, SFB 610). Funding for open access charge: Deutsche Forschungsgemeinschaft.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Sprinzl, M. and Cramer, F. (1979) The -C-C-A end of tRNA and its role in protein biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.*, **22**, 1–69.
2. Schimmel, P. and Yang, X.L. (2004) Two classes give lessons about CCA. *Nat. Struct. Mol. Biol.*, **11**, 807–808.
3. Weiner, A.M. (2004) tRNA maturation: RNA polymerization without a nucleic acid template. *Curr. Biol.*, **14**, R883–R885.
4. Vörtler, S. and Mörl, M. (2010) tRNA-nucleotidyltransferases: Highly unusual RNA polymerases with vital functions. *FEBS Lett.*, **584**, 297–302.
5. Betat, H., Rammelt, C. and Mörl, M. (2010) tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cell. Mol. Life Sci.*, February 14 [Epub ahead of print; doi: 10.1007/s00018-010-0271-4].
6. Yue, D., Weiner, A.M. and Maizels, N. (1998) The CCA-adding enzyme has a single active site. *J. Biol. Chem.*, **273**, 29693–29700.
7. Augustin, M.A., Reichert, A.S., Betat, H., Huber, R., Mörl, M. and Steegborn, C. (2003) Crystal Structure of the Human CCA-adding Enzyme: Insights into Template-independent Polymerization. *J. Mol. Biol.*, **328**, 985–994.
8. Li, F., Xiong, Y., Wang, J., Cho, H.D., Tomita, K., Weiner, A.M. and Steitz, T.A. (2002) Crystal Structures of the Bacillus stearothermophilus CCA-Adding Enzyme and Its Complexes with ATP or CTP. *Cell*, **111**, 815–824.
9. Okabe, M., Tomita, K., Ishitani, R., Ishii, R., Takeuchi, N., Arisaka, F., Nureki, O. and Yokoyama, S. (2003) Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J.*, **22**, 5918–5927.

10. Xiong, Y., Li, F., Wang, J., Weiner, A.M. and Steitz, T.A. (2003) Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol. Cell. Biochem.*, **12**, 1165–1172.
11. Toh, Y., Takeshita, D., Numata, T., Fukai, S., Nureki, O. and Tomita, K. (2009) Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. *EMBO J.*, **28**, 3353–3365.
12. Just, A., Butter, F., Trenkmann, M., Heitkam, T., Mörl, M. and 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.
13. Neuenfeldt, A., Just, A., Betat, H. and Mörl, M. (2008) Evolution of tRNA nucleotidyltransferases: A small deletion generated CC-adding enzymes. *Proc. Natl Acad. Sci. USA*, **105**, 7953–7958.
14. Zhu, L.Q., Cudny, H. and Deutscher, M.P. (1986) A mutation in *Escherichia coli* tRNA nucleotidyltransferase that affects only AMP incorporation is in a sequence often associated with nucleotide-binding proteins. *J. Biol. Chem.*, **261**, 14875–14877.
15. McGann, R.G. and Deutscher, M.P. (1980) Purification and characterization of a mutant tRNA nucleotidyltransferase. *Eur. J. Biochem.*, **106**, 321–328.
16. Bralley, P., Chang, S.A. and Jones, G.H. (2005) A phylogeny of bacterial RNA nucleotidyltransferases: *Bacillus halodurans* contains two tRNA nucleotidyltransferases. *J. Bacteriol.*, **187**, 5927–5936.
17. Tomita, K. and Weiner, A.M. (2001) Collaboration between CC- and A-adding enzymes to build and repair the 3'-terminal CCA of tRNA in *Aquifex aeolicus*. *Science*, **294**, 1334–1336.
18. Tomita, K. and Weiner, A.M. (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.
19. Tomita, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, N., Vassilyev, D.G. and Nureki, O. (2004) Structural basis for template-independent RNA polymerization. *Nature*, **430**, 700–704.
20. Betat, H., Rammelt, C., Martin, G. and Mörl, M. (2004) Exchange of regions between bacterial poly(A) polymerase and CCA adding enzyme generates altered specificities. *Mol. Cell*, **15**, 389–398.
21. Reichert, A.S., Thurlow, D.L. and Mörl, M. (2001) A eubacterial origin for the human tRNA nucleotidyltransferase? *Biol. Chem.*, **382**, 1431–1438.
22. Mörl, M., Lizano, E., Willkomm, D.K. and Hartmann, R.K. (2005) In Hartmann, R.K., Bindereif, A., Schön, A. and Westhof, E. (eds), *Handbook of RNA Biochemistry*, Vol. 1. Wiley-VCH, Weinheim, pp. 22–35.
23. Schürer, H., Lang, K., Schuster, J. and Mörl, M. (2002) A universal method to produce in vitro transcripts with homogeneous 3' ends. *Nucleic Acids Res.*, **30**, e56.
24. Cho, H.D., Verlinde, C.L. and Weiner, A.M. (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.
25. Tomita, K., Ishitani, R., Fukai, S. and Nureki, O. (2006) Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature*, **443**, 956–960.
26. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
27. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
28. Bryson, K., McGuffin, L.J., Marsden, R.L., Ward, J.J., Sodhi, J.S. and Jones, D.T. (2005) Protein structure prediction servers at University College London. *Nucleic Acids Res.*, **33**, W36–W38.
29. McGuffin, L.J., Bryson, K. and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404–405.
30. Rost, B., Yachdav, G. and Liu, J. (2004) The PredictProtein server. *Nucleic Acids Res.*, **32**, W321–W326.
31. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
32. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
33. Lindahl, E., Hess, B. and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.*, **7**, 306–317.
34. Kaminski, G., Friesner, R., Tirado-Rives, J. and Jorgensen, W. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*, **105**, 6474–6487.
35. Berendsen, H., Postma, J., Van Gunsteren, W., Dinola, A. and Haak, J. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.
36. Ryckaert, J., Ciccotti, G. and Berendsen, H. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.
37. Yue, D., Maizels, N. and Weiner, A.M. (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.
38. Raynal, L.C., Krisch, H.M. and Carpousis, A.J. (1998) The *Bacillus subtilis* nucleotidyltransferase is a tRNA CCA-adding enzyme. *J. Bacteriol.*, **180**, 6276–6282.
39. Sohlberg, B., Huang, J. and Cohen, S.N. (2003) The *Streptomyces coelicolor* polynucleotide phosphorylase homologue and not the putative poly(A) polymerase can polyadenylate RNA. *J. Bacteriol.*, **185**, 7273–7278.
40. Martin, G. and Keller, W. (2004) Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. *RNA*, **10**, 899–906.
41. Olsen, G.J., Woese, C.R. and Overbeek, R. (1994) The winds of (evolutionary) change: breathing new life into microbiology. *J. Bacteriol.*, **176**, 1–6.
42. Woese, C.R. (2002) On the evolution of cells. *Proc. Natl Acad. Sci. USA*, **99**, 8742–8747.
43. Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M. and Aksoy, S. (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat. Genet.*, **32**, 402–407.
44. Wren, B.W. (2002) Deciphering tsetse's secret partner. *Nat. Genet.*, **32**, 335–336.
45. Lizano, E., Scheibe, M., Rammelt, C., Betat, H. and 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.
46. Lin, G.C., Jaeger, J., Eckert, K.A. and Sweasy, J.B. (2009) Loop II of DNA polymerase beta is important for discrimination during substrate binding. *DNA Repair*, **8**, 182–189.
47. Juarez, R., Ruiz, J.F., Nick McElhinny, S.A., Ramsden, D. and Blanco, L. (2006) A specific loop in human DNA polymerase mu allows switching between creative and DNA-instructed synthesis. *Nucleic Acids Res.*, **34**, 4572–4582.
48. Balendiran, G.K., Molina, J.A., Xu, Y., Torres-Martinez, J., Stevens, R., Focia, P.J., Eakin, A.E., Sacchettini, J.C. and Craig, S.P. III (1999) Ternary complex structure of human HGPRase, PRPP, Mg<sup>2+</sup>, and the inhibitor HPP reveal the involvement of the flexible loop in substrate binding. *Protein Sci.*, **8**, 1023–1031.
49. Greenwald, J., Le, V., Butler, S.L., Bushman, F.D. and Choe, S. (1999) The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry*, **38**, 8892–8898.
50. Ko, T.P., Chen, Y.K., Robinson, H., Tsai, P.C., Gao, Y.G., Chen, A.P., Wang, A.H. and Liang, P.H. (2001) Mechanism of product chain length determination and the role of a flexible loop in *Escherichia coli* undecaprenyl-pyrophosphate synthase catalysis. *J. Biol. Chem.*, **276**, 47474–47482.
51. Joseph, E., Wei, W., Tittmann, K. and Jordan, F. (2006) Function of a conserved loop of the beta-domain, not involved in thiamin diphosphate binding, in catalysis and substrate activation in yeast pyruvate decarboxylase. *Biochemistry*, **45**, 13517–13527.