Published in final edited form as:

Biochim Biophys Acta. 2008 April ; 1779(4): 206–216. doi:10.1016/j.bbagr.2007.12.003.

Determinants of substrate specificity in RNA-dependent nucleotidyl transferases

Georges Martin^a, Sylvie Doublé^b, and Walter Keller^{a,*}

^a Department of Cell Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland ^b Department of Microbiology and Molecular Genetics, University of Vermont, Given Building E314-A, Burlington, VT 05405 USA

Abstract

Poly(A) polymerases were identified almost fifty years ago as enzymes that add multiple AMP residues to the 3' ends of primer RNAs without use of a template from ATP as cosubstrate and with release of pyrophosphate. Based on sequence homology of a signature motif in the catalytic domain, poly(A) polymerases were later found to belong to a superfamily of nucleotidyl transferases acting on a very diverse array of substrates. Enzymes belonging to the superfamily can add from single nucleotides of AMP, CMP or UMP to RNA, antibiotics and proteins but also homopolymers of many hundred residues to the 3' ends of RNA molecules.

The recently reported structures of several nucleotidyl transferases facilitate the study of the catalytic mechanisms of these very diverse enzymes. Numerous structures of CCA-adding enzymes have now revealed all steps in the formation of a CCA tail at the 3' end of tRNAs. In addition, structures of poly(A) polymerases and uridylyl transferases are now available as binary and ternary complexes with incoming nucleotide and RNA primer. Some of these proteins undergo significant conformational changes after substrate binding. This is proposed to be an indication for an induced fit mechanism that drives substrate selection and leads to catalysis. Insights from recent structures of ternary complexes indicate an important role for the primer molecule in selecting the incoming nucleotide.

Keywords

nucleotidyl transferase; terminal uridylyl transferase; CCA-adding enzyme; poly(A) polymerase; catalytic mechanism

1. Introduction

Based on sequence homology with the DNA polymerase β (Pol β) signature helix-turn motif hG[G/S]_{x9-13}Dh[D/E]h (x, any; h, hydrophobic amino acids) a superfamily of nucleotidyl transferases (NTrs) consisting of enzymes with a surprisingly large substrate diversity has been identified [1-4]. Pol β is mainly involved in base-excision repair and in DNA double-strand-break repair and homologues are present in all three kingdoms of life (reviewed in [5]). A

*Corresponding author: Tel.: +41 61 267 2060, Fax: +41 61 267 2079, E-mail address: E-mail: walter.keller@unibas.ch.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

related enzyme, terminal deoxynucleotidyl transferase (TdT), takes part in generating the antibody diversity in higher eukaryotes by random dNTP addition during V(D)J recombination [6]. However, most members of the superfamily transfer ribonucleoside triphosphates instead of dNTPs to the recipient substrates and do not use a nucleic acid template for nucleotide selection. These ribonucleotidyl transferases (rNTrs) are ubiquitous enzymes and carry out many different functions. The ancestors of NTrs are the minimal nucleotidyl transferases (MNTs), proteins of less than 100 amino acids found in some archaea and eubacterial species. MNTs consist of only the core catalytic domain, sometimes associated with a second subunit, which could specify substrate binding (Fig. 1A) [7]. Other members of the rNTr subfamily were recently described and characterized biochemically. Crystal structures were obtained for a subset of these enzymes, namely eukaryotic poly(A) polymerases (PAPs; Fig. 1B), terminal uridylyl transferases (TUTases; Fig. 1C), CCA adding enzymes (CCATrs; Fig. 1D), protein nucleotidyl transferases, 2' 5' oligo(A) synthetases (OASs) and antibiotics nucleotidyl transferases (reviewed in [8]). Some rNTrs transfer a single nucleotide to the 3' hydroxyl end of a primer and others, in particular in cases where the recipient substrate is an RNA primer, add from a few to several hundred nucleotides resulting in homopolymeric poly(A) or poly(U) tails. In eukaryotes these tails have been found to be highly homogeneous i.e. few other nucleotides are incorporated. In contrast, the error rate during poly(A) synthesis in eubacteria was found to be higher [9]. The nucleotide substrate in most rNTrs consists of one specific ribonucleotide (ATP or UTP) with the exception of CCATrs, which synthesize the trinucleotide sequence CCA at the 3' ends of tRNA precursors. Poly(A) tails added to pre-mRNAs by canonical PAPs function to stabilize mRNAs by preventing premature degradation and are required for the transport of the RNAs to the cytoplasm and for their efficient translation into proteins. In contrast, poly(A) and poly(U) tails generated by non-canonical PAPs and TUTases in eucaryotes, as well as poly(A) tails found in eubacterial RNAs are involved in RNA turnover [10] (reviewed in [8,11]).

The increasing number of rNTrs made it necessary to update their classification. rNTrs are currently classified into canonical and noncanonical rNTrs [8,12], according to their functional and structural attributes. The canonical rNTrs include the classical PAPs first described in eukaryotes, which polyadenylate pre-mRNAs in the nucleus and in some cases, in the cytoplasm [13-15]. Depending on the organism one (yeast and *Drosophila melanogaster*) to three genes coding for PAP have been found in vertebrates and in *C. elegans*. Non-canonical rNTrs can be distinguished from canonical rNTrs mainly by a characteristic nucleotide recognition motif (NRM) which seems to have evolved independently of the NRM in canonical NTrs [1,8]. The signature motif defining NRM 1 is MPIITPAYPQQN, which is almost invariant in canonical PAPs of all eukaryotes. The consensus for the NRM 2 motif of non-canonical PAPs is hx[I/L/V][E/Q][E/D/N]Phx₄N (x, any; h, hydrophobic amino acids) with only P-6 and N-12 being conserved between the two motifs.

Among the recently described non-canonical rNTrs are the cytoplasmic PAP GLD-2 (germ line development 2) of *C. elegans*, which reactivates maternal mRNAs to undergo translation in early development [16] and the yeast PAPs Trf4 and Trf5, which are involved in quality control and turnover of misfolded RNAs [17-20]. Based on the presence of an NRM-2 motif, TUTases and the recently described poly(U) polymerases (PUPs) [21,22] can also be assigned to the non-canonical rNTr subfamily, which thus comprises both PAPs and TUTases. The distinction between TUTases and PUPs is at the moment unclear because protist TUTases have been described that can also synthesize long poly(U) tails [21,22]. Moreover, because of the lack of structural information it is still not known how non-canonical PAPs select ATP since they contain an NRM of type 2 that is almost identical to that of TUTases.

Furthermore, NTrs have been classified according to two submotifs in the catalytic site helical-turn motif described above [4]. This seemingly small difference has proven to be of significant

help for establishing the phylogeny of NTrs (Fig. 2). The division into class I and class II NTrs is based on additional amino acids in the region of the helical-turn motif, which results in a helix with an extra turn in class II NTrs. Class II NTrs include bacterial PAPs and CCATrs and eukaryotic CCATrs (see below). In contrast, archaeal CCATrs belong to class I rNTrs (Fig. 2) and show strong structural homology to canonical and non-canonical rNTrs (Fig. 1) [1]. It is also possible that the latter are descendents of an archaeal enzyme. This could explain the close similarity in substrate selection and in the catalytic mechanism of these two groups of rNTrs.

2. Structure of active sites in U-adding enzymes

TUTases were first described in trypanosomal mitochondria and are involved in uridine insertion/deletion RNA editing. RNA editing TUTase 1 (RET1) is a processive enzyme and is required for the uridylation of guide RNAs [23] and is also involved in mitochondrial RNA turnover [24]. RNA editing TUTase 2 (RET 2) is a subunit of a multi-protein editing complex (or 20 S editosome) and participates in guide-RNA-dependent U-insertion into mitochondrial pre-mRNAs [25]. TUTases have also been found in plants [26] and a human U6 snRNA-specific TUTase was recently isolated, which adds four nontemplated U residues to the 3' end of the U6 snRNA precursor [27].

All TUTases are structurally related and share a NTr signature as well as a NRM of type 2 reminiscent of non-canonical rNTrs. Comparison of active site architectures in recently published structures of TUTases reveals a high degree of specificity for UTP binding. Fig. 3A depicts the example of the *T. brucei* TUT4 (*TbTUT4*) active site with bound UTP [28] that is very similar to the RET2 active site [29]. The triphosphate moiety in the *TbTUT4* structure contributes three non-bridging phosphate oxygens to the coordination of Mg^{2+} , whereas two carbonyl oxygens of Asp66 and Asp68 and an additional water complete the metal coordination shell. Two γ -phosphate oxygens interact with the Ser65, Lys169, Lys173 and Ser188 side chains. The 2' and 3' hydroxyls of the ribose sugar are contacted via hydrogen bonds by the Ser148 hydroxyl and the Asn147 carbonyl and these interactions are thought to prevent the binding of deoxynucleotides. The uracil base is held in place by several interactions with nearby water molecules, which form an extensive network of hydrogen bonds. These water molecules are forming hydrogen bonds with Arg307, Asp297, Glu300, and Arg141 and interact with the uracil base via specific hydrogen bonds with the O2, N3 and O4 positions. N3 and O4 correspond to the Watson-Crick face of thymine, which hydrogen bonds with adenine in template-dependent DNA polymerases.

The fact that non-canonical PAPs and TUTases share an NRM 2 motif would predict that they select the same nucleotide. Coincidentally the *Schizosaccharomyces pombe* non-canonical rNTr Cid1 (for "Caffeine induced death 1"), was initially found to have polyadenylating activity [30]. However, when the same recombinant protein was expressed and purified as a native multiprotein complex from *S. pombe* instead of *E. coli* it only exhibited poly(U) adding activity [22]. This suggests that the non-canonical rNTrs may only need a minor modification or a small conformational change to switch between A and U selection. This is surprising because adenine and uracil differ with respect to size and Watson-Crick interface. Moreover, the recently reported crystal structures of *TbTUT4* in complex with the nucleotides ATP, GTP and CTP in addition to UTP indicate that nucleotide binding in these enzymes is rather promiscuous and that specificity must be obtained by additional means [31]. A structure of *TbTUT4* with nucleotide and UMP (an RNA mimic) or the postreaction complex with the UpU dinucleotide (Fig. 3B) together with biochemical tests of different types of primers suggests that a U residue is preferred at the penultimate position of the RNA substrate and that the uracil base at the 3' terminus of the RNA primer is involved in correct U addition [31]. The authors propose a dual role for the RNA substrate: first, the ultimate uracil base at the primer's 3' end forms a stacking interaction with the uracil base of the bound nucleotide, which also stacks

with an essential tyrosine and second, that the primer contributes to the formation of a productive catalytic metal binding site. The results also suggest that the penultimate U of the primer is in fact bound via hydrogen bonds to Arg126, which is a U specific interaction. This could be a control function to ensure error free poly(U) synthesis and it was demonstrated that mutation of Arg126 results in a completely inactive enzyme because of loss of RNA binding but not UTP binding [28].

3. Structure of active sites in CCA-adding enzymes

CCAtrs catalyze the post-transcriptional addition of a CCA trinucleotide to the 3' end of pre-tRNAs. The CCA end is essential in all organisms and is encoded only in the tRNA genes of some eubacterial species where CCAtrs repair missing or defective CCA ends. In all other organisms CCA is added *de novo* after tRNA processing where it becomes the acceptor for the corresponding amino acid. The mechanism for CCA addition and in particular how C, C and A are sequentially selected and added to the tRNA substrate remained a mystery for a long time. The first crystal structures of CCAtrs of human, archaeal (*Archaeoglobus fulgidus*) and eubacterial (*Bacillus stearothermophilus*) sources were determined in the apo form or as binary complexes with CTP or ATP [32-34]. Specific binding of the nucleotide bases cytosine or adenine were found in the *B. stearothermophilus* CCAtr, where the Watson-Crick face of cytosine formed two H-bonds with the guanidinium group of Arg157 and one H-bond with the Asp154 carboxylate [33]. ATP, on the other hand, was selected via two H-bond interactions between the adenine base and the same two amino acids except that in this case Arg157 was repositioned to allow only one hydrogen bond. This selection mechanism excludes GTP and UTP from binding because their pattern of hydrogen donors and acceptors are the reverse of those of ATP and CTP. These interactions are identical to those seen in C:G and A:U Watson-Crick base pairing. The active site pocket in eubacterial CCAtrs is providing a protein template and allows the specific selection of the correct nucleotides for CCA synthesis. In contrast, binding of the nucleotides appeared nonspecific in the absence of the tRNA substrate in the crystal structure of the archaeal *A. fulgidus* CCAtr [34]. However, in a recent crystal structure of a ternary complex that included an RNA minihelix mimicking a tRNA, the *A. fulgidus* CCAtr only bound CTP and ATP and excluded GTP and UTP by specific hydrogen bonding interactions of the nucleotides with Arg224 and backbone phosphates of the tRNA (Fig. 3C) [35]. These specific interactions occur in conjunction with a sequential rearrangement of the binding pocket to accommodate the growing 3' end and the incoming nucleotide and thus ensure the correct synthesis of CCA synthesis [35,36]. Recently reported ternary structures of PAPs and TUTases with nucleotide and primer similarly suggest that specificity for nucleotide selection results from interactions of the incoming nucleotide with active site amino acids as well as with the RNA primer [31,37].

Part of the ribonucleoprotein template pocket consists of a conserved β -turn structure which is structurally and functionally conserved in all CCAtrs [38]. A variant of this structure is also found in the bacterial PAPs where it is most likely involved in primer recognition and thus far is the only motif found to distinguish class II CCAtrs and PAPs [39]. Moreover, there is still no convincing explanation for the ability of the class I CCA adding enzymes to switch from C to A after C75 addition because any pocket large enough to accommodate A is expected to also fit the smaller C. Additional structural data are needed to unveil the details of this fascinating mechanism.

4. Structure of active sites in canonical poly(A) polymerases

Mammals contain three canonical PAPs called PAP- α , - β and - γ that are encoded by different genes. The first reported structures were the yeast and mammalian canonical PAPs in complex with 3' dATP [40,41]. The two polymerases adopt the same overall fold with a catalytic, a

central and an RNA binding domain (RBD). Mammalian and yeast PAP are about 48% identical in their protein sequence within the catalytic and central domain and have substantially less sequence conservation between the RNA binding domains. In both structures the triphosphate moiety of the nucleotide is seen bound in a similar fashion with coordination of the phosphates via non-bridging oxygens by divalent metal ions and additional contacts with amino acids from the central domain. In the bovine PAP- α structure the base of the Mn-3'dATP nucleotide was bound specifically but positioned in a nonproductive orientation, which would not allow proper alignment of the primer 3' OH with the α phosphate of the incoming nucleotide for nucleophilic attack [41]. In the yeast PAP structure on the other hand, there were no adenine-specific interactions with the protein but the base was stacked on the base of a second ATP molecule that was proposed to mimic the primer 3' end [40].

In a subsequent bovine PAP structure complexed with 3'dATP in the presence of Mg^{2+} instead of Mn^{2+} the adenine base was found in a different orientation and interacted with water molecules and residues at the bottom of the central domain (Fig. 3E) [42]. The stacking interaction between Val247 and the adenine base could contribute to the specificity for ATP by selecting against the smaller pyrimidines cytosine and uracil. Other rNTrs, such as the TUTase TUT4 or vaccinia PAP (Fig. 3A,D) use similar stacking interactions with side chains of Met or Tyr to orient nucleotide bases in their active site [28,43].

In the bovine PAP structure with Mg^{2+} residues Asn202 and Thr317 interact with the adenine base via weak hydrogen bonds [44]. Site directed mutation of the two residues Asn202 and Thr317 produced kinetic parameters (K_M) that confirm an involvement of these residues in nucleotide binding (Fig. 3E) [42]. In addition, the natural substrate ATP (instead of 3'dATP used for crystallization) may form additional hydrogen bonding interactions with its 3' OH and could result in a different type of hydrogen-bonded water network. In contrast, the water (Wat2043) bridge between the N6 amine of the adenine (hydrogen donor) and the backbone oxygen (hydrogen acceptor) of Ser225 is specific and is unlikely to form with a guanine base, although additional water interactions could form with a bound GTP. Furthermore, primer binding and closing of the catalytic and central domains could lead to rearrangements within the active site resulting in more stable contacts of residues close to the base, such as N202 and T317 (see below; [45]). The sum of the interactions with the adenine seen in this binary complex can be considered as specific for purines, which therefore would also allow GTP binding. However, when bovine PAP was tested in the presence of Mn^{2+} and 0.5 mM nucleotide GTP incorporation was not detectable and UTP was added to primer at 0.2% and CTP at 0.5% compared to ATP [15].

The two bovine PAP structures solved with Mn^{2+} and Mg^{2+} differed by about 2.5° in the orientation of the RNA binding domain relative to the catalytic domain [41,42]. In addition, recently reported crystal forms of yeast PAP revealed structures with varying open or closed states with respect to the catalytic cleft. The N-terminal and C-terminal domains were found to move independently as rigid bodies along two defined axes of rotation [45]. In addition, steady state fluorescence and acrylamide quenching experiments with 2-aminopurine substrates suggested that the closed domain confirmation was stabilized upon recognition of the correct substrate (see below) [45]. Taken together these results indicate that PAPs are inherently flexible enzymes and suggest a closure of the two domains to occur upon substrate binding as seen in most DNA polymerases, such as DNA polymerase β [46,47]. Pol β belongs to the pol X family of polymerases which itself is part of the same superfamily of nucleotidyl transferases as PAP [1]. Pol β was shown to undergo a large conformational change upon binding of DNA and the correct incoming nucleotide [47]. We note however that a large-scale domain movement is not universal among Pol X DNA polymerases: Pol λ does not seem to undergo a conformational change during catalysis. What was observed instead when

comparing binary and ternary complex structures was a repositioning of key side chains and the template strand (reviewed in [48]).

A recently reported ternary structure of the yeast PAP mutant D154A with MgATP and a short (A)₅ primer provides new insights into the mechanism of poly(A) addition [37]. The mutation of Asp154 (an aspartic acid needed for catalysis) made PAP catalytically inactive without compromising substrate binding except that only one of the catalytic metal ions was bound. In this case, an inactive form of the enzyme allowed the visualization of a stalled ternary complex. In contrast to earlier binary structures the enzyme has undergone significant domain movements and is now found in a closed conformation with the N-terminal catalytic domain moved in the direction of the C-terminal RBD forming a tunnel that encloses MgATP and the RNA primer in addition to about 30 water molecules (Fig. 3F). Adenine bases at primer positions -1 to -3 (-1 is the A at the 3' end) are completely buried and the 2' hydroxyl groups at the -1 to -3 positions are directly hydrogen bonded to the enzyme, which would disfavor binding of a DNA substrate. Primer position -4 is located at the surface of the protein and is partially exposed to the solvent and position -5 is bound by another symmetry related PAP molecule and it is unclear whether that could affect the positions of the primer bases -1 through -4. As in the earlier yeast PAP structure bound to ATP [40], the adenine base of the incoming nucleotide is not specifically recognized and is stacked against the last adenylyate of the oligo (A) primer. The authors propose a combination of several indirect factors to explain specificity for A addition. Part of the specificity could result from π -stacking interactions between the primer's -1 adenine and the base of the incoming nucleotide, thereby favoring purine over pyrimidine bases. Upstream pre-mRNA cleavage products mostly contain an A at the 3' end and subsequent additions intrinsically are As. In addition, residues -2 and -3 are also engaging in stacking interactions with aromatic residues or hydrophobic moieties of amino acid side chains (Fig. 3F). Unexpectedly, the bases of the primer in this ternary structure are bound to the active site in a completely unfolded manner, with angles of between roughly 110 and 175 degrees among the -1 to -3 positions (Fig. 3F) [37]. The authors postulate that the primer dissociates from the active site after each addition even under cellular conditions of locally high poly(A) concentrations as is the case in processive polyadenylation. This should be facilitated by the large number of buried waters in the closed complex which should lower the kinetic barrier to domain movements and minimize the extent of substrate desolvation [37]. It remains to be clarified whether the three bases close to the 3' end during processive polyadenylation switch to binding in an energetically more favorable way or become π -stacked to allow fast translocation.

Primer translocation has not been studied extensively in rNTrs except in vaccinia PAP [49]. Primer/template translocation studies in T7 RNA polymerase [50] could also apply to the processive PAPs and TUTases. In this report, it was postulated that the dissociation of the product PP_i after catalysis would lead to a protein conformational change resulting in primer translocation.

In the above examples domain movements seem to be required for polynucleotide synthesis, yet in some rNTrs catalysis can occur in the absence of domain motion. In cells infected by poxviruses, polyadenylation of viral mRNAs is carried out by a poly(A) polymerase heterodimer, which is composed of the catalytic subunit VP55 and a processivity factor VP39. The structure of the polymerase was determined in the apo form and as a binary complex with ATP- γ -S [43]. Residues of the central domain within the active site of VP55 make specific interactions with the adenine of the ATP analog (Fig. 3D). A model for RNA primer binding that involves VP55 and VP39 was proposed, suggesting that VP39 functions as a processivity factor by partially enclosing the RNA primer at the heterodimer interface. Yet, there is no indication of flexibility between domains as in the yeast and mammalian PAP structures. VP55

forms a compact globular structure and does not consist of three flexible domains as seen in eukaryotic canonical PAPs (see below and Table 1).

5. Catalytic mechanism in poly(A) polymerases

Catalytic sites in NTrs typically consist of a two-layer $\alpha\beta$ fold composed of a five stranded β -sheet backed by two α -helices. Located on top of the β -sheet are three Asp residues (D113, D115 and D167 in bovine or human PAP- α), which coordinate two Mg^{2+} ions required for catalysis. One of the metals enters the site with the nucleotide ($MgATP^{2-}$) and leaves as $MgPP_i^{2-}$ after catalysis. The other metal coordinates the α -phosphate of the nucleotide with the 3' OH of the primer and stabilizes the negative charge that develops during the transition state. NTrs of the Pol β superfamily act according to this two-metal ion catalytic mechanism like all metal-dependent DNA or RNA polymerases [51]. Fig. 4 depicts the active site structure of Pol β in which the complete coordination shells for two metal ions are visible [46]. In all presently known rNTr structures the coordination shells are incomplete.

Steady state kinetic analysis has described the PAP reaction ($A_n + ATP \rightleftharpoons A_{n+1} + PP_i$) as a rapid equilibrium random mechanism [52]. This indicates that in an *in vitro* reaction in the absence of assisting processivity factors the order of substrate binding is random and the products leave the active site after catalysis. Inhibition studies showed that substrate dissociation is fast relative to the internal conversion step, consistent with a rapid equilibrium model [52]. In contrast, PAP is expected to employ a non-random or ordered mechanism of substrate assembly within the active site *in vivo* [42]. The product A_{n+1} would - after translocation from the nucleotide site to the primer site - instantly become the new primer and therefore be present in the active site before the next nucleotide enters. This is possible because the primer product is prevented from dissociation by processivity factors such as the cleavage and polyadenylation factor CPSF and the poly(A) binding protein PABPN1 in mammals [53] or the cleavage and polyadenylation factors CF I and CPF in yeast [54].

Studies of the sulfur elemental effect of [α -thio] substituted ATP and CTP and additional kinetics tests suggest that ATP utilizes a ground-state destabilization mechanism for rate acceleration and CTP does not [52]. In addition, it was proposed that an induced fit mechanism is responsible for the nucleotide specificity that is first mediated by stabilization of both the transition state and ground state by using energy derived from the recognition of the adenine base of ATP. This binding energy is then utilized to increase the free energy of the ground state (ground state destabilization) to promote rate acceleration [52].

Results of steady state fluorescence experiments of yeast PAP suggested that an enzyme-substrate ternary complex is stabilized in a more closed conformational state [45]. Moreover, comparison of yeast PAP structures of different crystal forms indicates that the N-terminal and C-terminal domains of PAP move independently as rigid bodies along two well defined axes of rotation [45]. Taken together, these results suggest that substrates and products can rapidly associate or dissociate from an open conformation of PAP. After recognition of the correct substrates in the forward direction of the reaction a closed enzyme conformation is stabilized and efficient catalysis is made possible.

6. Primer substrate recognition and processivity

All nucleotidyl transferases require a recipient or primer substrate to which a nucleoside monophosphate can be attached. Most NTrs use a double stranded DNA or single stranded RNA primer. Exceptions are the 2'-5' oligo(A) synthetases that produce oligo(A) with 2'-5' phosphodiester bonds in sizes of up to 30 nucleotides by starting with an ATP as primer (reviewed in [55]). Other exceptions are antibiotics resistance factors such as kanamycin nucleotidyl transferase, which adds AMP to aminoglycoside antibiotics and inactivates

kanamycin by adenylation (reviewed in [56]). Nucleotides can also be transferred to proteins by Pol β type NTrs. For example, the P_{II} uridylyl transferase (GlnD) in bacteria modifies the P_{II} regulatory protein (GlnB) by uridylylation and deuridylylation [57].

Recognition of non-nucleotide primer substrates such as aminoglycosides or protein P_{II} most likely works via interaction with specific ligands followed by release of the products after nucleotide addition in a fashion similar to distributive nucleotidyl transfer to RNA primers. The molecular mechanism of processive polynucleotide polymerases is very different because the primer DNA or RNA translocates after each nucleotide addition.

Recognition and translocation of RNA primers in nontemplated rNTrs have been studied in vaccinia PAP [49] and in canonical PAPs [3,58,59]. Poxvirus PAP (VP55) recognizes stretches of oligo(U) close to the 3' end of the primer RNA and then adds processively ~25-30 AMP residues to the 3' end [49]. The RNA path on the VP55/VP39 heterodimer leads from the active site along a groove between the two subunits and ends at the opposite side of the dimer complex. VP39 is a subunit with cap-specific 2'O-methyltransferase activity on the 5' ends of viral mRNAs and VP39 is proposed to be also a processivity factor for VP55 because of its primer binding activity (Fig. 5A) [43].

Biochemical experiments with *S. cerevisiae* PAP have shown that three ribonucleotides at the 3' end of the primer are sufficient for recognition such that the remainder of the primer can consist of DNA, indicating that binding after these three nucleotides is only specific for the phosphate backbone [60]. In particular, Phe153 located on a β -turn in bovine PAP seems to be essential for binding to the RNA primer close to the 3' end [42] and the corresponding Phe140 in yeast PAP was recently shown to contact the third adenylate residue of the RNA primer [37]. A second site of interaction with the primer in bovine PAP lies on the outer side of the RBD where Arg or Lys residues belonging to a nuclear localization signal motif are responsible for interactions with nucleotides about eight to ten bases away from the 3' end of the RNA [3]. A similar second binding site was also found in yeast PAP [60]. It should be noted that these conclusions are derived from *in vitro* experiments and the situation is likely to be different *in vivo* where the emerging poly(A) tail beginning at residue -4 is expected to be handed over to processivity factors.

Class I bacterial and human CCA adding enzymes differ from the highly homologous bacterial PAPs in the way they bind the substrate RNA. Both enzymes contain β -turn structures close to the active site and sequence motifs located on these β -turns seem to be specific for CCA or poly(A) addition [38,39]. It was also demonstrated that in CCAtr the tRNA precursor substrate does not translocate during CCA synthesis: If the tRNA was cross-linked to the enzyme CCA was still faithfully added to the tRNA's 3' end [61]. In addition, experiments with the *E. coli* CCAtr and PAP have shown that when the N-terminal catalytic and C-terminal RNA binding domains were exchanged between the two proteins, the PAP C-terminus allowed translocation of an RNA primer, thus resulting in a poly(CCA) tail added to tRNA when combined with the catalytic domain of CCAtr [62]. Conversely, when the C-terminal RNA binding domain of CCAtr was fused to the catalytic domain of bacterial PAP, the chimera acquired a CCA adding activity and added a single CCA. This unexpected result indicated that the PAP catalytic domain can be forced to synthesize CCA instead of its natural product poly(A) when the tRNA substrate is prevented from translocating. This cryptic activity could originate from an interconversion of these two proteins during evolution [4].

Besides these direct binding interactions between RNA primers and polymerases, several protein factors were described that act to increase the processivity of the catalytic reaction of these enzymes. By definition, a processivity factor acts by preventing the dissociation of the primer from the enzyme after each nucleotide addition. This way, the reaction rate can become

very high and can be in the order of several hundred additions per second. Polymerisation without processivity factors is defined as distributive.

Bovine PAP was shown to engage with several proteins that act as processivity factors. Both the 160 kDa subunit of cleavage and polyadenylation specificity factor (CPSF-160) and the factor interacting with PAP 1 (Fip1) are part of the multiprotein CPSF complex, which first cleaves the mRNA precursor at the poly(A) site near the conserved AAUAAA motif [63]. After cleavage, these two subunits interact directly with PAP and stimulate its catalytic activity slightly by tethering the RNA to the polymerase [64]. Binding of nuclear poly(A) binding protein 1 (PABPN1) to the polyadenylation complex confers high processivity to the reaction and results in the synthesis of a poly(A) tail of ~250 nucleotides without dissociation of the complex. After addition of the poly(A) tail the reaction becomes distributive and then terminates [65]. Fip1 was also shown to contribute substrate specific binding in the CPSF complex by enhancing the affinity to U-rich elements near the cleavage site of the pre-mRNA and to stimulate polyadenylation *in vitro* [64]. In contrast, the yeast Fip1 homologue reduced polyadenylation activity of yeast PAP when assayed *in vitro*, suggesting that Fip1 is a regulatory factor rather than a factor that exclusively stimulates PAP activity (Fig. 5B) [66].

Substrate and sequence specificity is often contributed to rNTrs by subunits of a multiprotein complex. A good example is the RNA editing TUTase RET2, which is one of the many subunits of the 20 S editosome where its task is to add single uridylate residues to the 3' ends of cleaved editing intermediates [25]. Furthermore, the non-canonical PAP GLD-2 of the nematode *C. elegans*, which only consists of a catalytic and a central domain, associates with GLD-3, a homolog of Bicaudal-C, which contributes specific RNA binding to GLD-2 (Fig. 5C) [16]. It has also been shown that the GLD-2/GLD-3 complex directly targets the *gld-1* pre-mRNA [67]. The GLD-1 protein is a STAR/Quaking translational repressor that promotes entry into the meiotic cell cycle.

Other recently described non-canonical PAPs are the yeast Trf4 and Trf5 proteins. They were found to participate in RNA surveillance and turnover pathways that control the levels of a large array of nontranslated RNAs in *S. cerevisiae* [17-20]. Trf4 in complex with the cofactors Air1 or Air2 is able to distinguish among many different types of RNAs and can also detect hypomethylated tRNAs that are assumed to be misfolded (Fig. 5D). The Trf4 or TRAMP (for Trf4/Air2/Mtr4 polyadenylation) complex then promotes the complete degradation of these defective RNAs in the exosome by repeated polyadenylation of the degradation intermediates [17,19]. The poly(A) tails of the target RNAs and their degradation intermediates serve as binding platforms for the exosome. The exosome is a multiprotein complex involved in RNA turnover for which a structure was recently reported [68] (reviewed in [69]).

Table 1 summarizes features found in a selected set of rNTrs such as processivity, flexibility of the protein domains, the number of nucleotides added and whether the protein has an induced fit capability and contains specificity for substrate selection. The features "Flexibility" and "Induced fit" are most likely connected because the induced fit mechanism is defined as rearrangement of protein domains after substrate binding. Information on nucleotide specificity is based on structural data of binary and ternary complexes. In light of recent ternary structures, specificity for the incoming nucleotide can also be contributed by the primer [31,37]. Many of the characteristics listed in Table 1 are not precisely known and necessitate further investigation.

7. Conclusions and outlook

Although a wealth of information has been obtained recently by investigating the very diverse family of nucleotidyl transferases, many questions are still open. Much is known about

nucleotide recognition: in some rNTrs the incoming nucleotide was found to be recognized via specific interactions with active site residues, even in the absence of a primer substrate. In other structures of mostly processive rNTrs, binding of the nucleotide appears to be only partially selective. Recently reported structures of *TbTUT4* and yeast PAP in complex with nucleotide and primer RNA have now widened our understanding of substrate recognition and nucleotide selection. It appears that selection of the correct nucleotide is in part determined by the primer 3' end where stacking interactions between nucleotide bases play an important role. Future investigations should focus on the molecular mechanisms of catalysis in a fully assembled active site and on primer translocation in the context of processive polynucleotide synthesis.

Acknowledgements

This work has been supported by the University of Basel and the Swiss National Science Fund. SD acknowledges support from the NIH (Grant GM62239).

References

1. Aravind L, Koonin EV. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res* 1999;27:1609–1618. [PubMed: 10075991]
2. Holm L, Sander C. DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem Sci* 1995;20:345–347. [PubMed: 7482698]
3. Martin G, Keller W. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. *EMBO J* 1996;15:2593–2603. [PubMed: 8665867]
4. Yue D, Maizels N, Weiner AM. 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* 1996;2:895–908. [PubMed: 8809016]
5. Ramadan K, Shevelev I, Hübscher U. The DNA-polymerase-X family: controllers of DNA quality? *Nat Rev Mol Cell Biol* 2004;5:1038–1043. [PubMed: 15573140]
6. Thai TH, Kearney JF. Isoforms of terminal deoxynucleotidyltransferase: developmental aspects and function. *Adv Immunol* 2005;86:113–136. [PubMed: 15705420]
7. Lehmann C, Pullalarevu S, Krajewski W, Willis MA, Galkin A, Howard A, Herzberg O. Structure of HI0073 from *Haemophilus influenzae*, the nucleotide-binding domain of a two-protein nucleotidyl transferase. *Proteins* 2005;60:807–811. [PubMed: 16028221]
8. Martin G, Keller W. RNA-specific ribonucleotidyl transferases. *RNA* 2007;13:1834–1849. [PubMed: 17872511]
9. Yehudai-Resheff S, Schuster G. Characterization of the *E.coli* poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence. *Nucleic Acids Res* 2000;28:1139–1144. [PubMed: 10666455]
10. Shen B, Goodman HM. Uridine addition after microRNA-directed cleavage. *Science* 2004;306:997. [PubMed: 15528436]
11. Hall TM. Poly(A) tail synthesis and regulation: recent structural insights. *Curr Opin Struct Biol* 2002;12:82–88. [PubMed: 11839494]
12. Stevenson AL, Norbury CJ. The Cid1 family of non-canonical poly(A) polymerases. *Yeast* 2006;23:991–1000. [PubMed: 17072891]
13. Lingner J, Kellermann J, Keller W. Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. *Nature* 1991;354:496–498. [PubMed: 1840648]
14. Raabe T, Murthy KG, Manley JL. Poly(A) polymerase contains multiple functional domains. *Mol Cell Biol* 1994;14:2946–2957. [PubMed: 8164653]
15. Wahle E. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J Biol Chem* 1991;266:3131–3139. [PubMed: 1993684]

16. Wang L, Eckmann CR, Kadyk LC, Wickens M, Kimble J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* 2002;419:312–316. [PubMed: 12239571]
17. Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes Dev* 2004;18:1227–1240. [PubMed: 15145828]
18. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* 2005;121:713–724. [PubMed: 15935758]
19. Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W. A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol* 2005;3:e189. [PubMed: 15828860]
20. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* 2005;121:725–737. [PubMed: 15935759]
21. Kwak JE, Wickens M. A family of poly(U) polymerases. *RNA* 2007;13:860–867. [PubMed: 17449726]
22. Rissland OS, Mikulasova A, Norbury CJ. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol Cell Biol* 2007;27:3612–3624. [PubMed: 17353264]
23. Aphasizhev R, Sbicego S, Peris M, Jang SH, Aphasizheva I, Simpson AM, Rivlin A, Simpson L. Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell* 2002;108:637–648. [PubMed: 11893335]
24. Ryan CM, Read LK. UTP-dependent turnover of *Trypanosoma brucei* mitochondrial mRNA requires UTP polymerization and involves the RET1 TUTase. *RNA* 2005;11:763–773. [PubMed: 15811918]
25. McManus MT, Adler BK, Pollard VW, Hajduk SL. *Trypanosoma brucei* guide RNA poly(U) tail formation is stabilized by cognate mRNA. *Mol Cell Biol* 2000;20:883–891. [PubMed: 10629045]
26. Zabel P, Dorssers L, Wernars K, Van Kammen A. Terminal uridylyl transferase of *Vigna unguiculata*: purification and characterization of an enzyme catalyzing the addition of a single UMP residue to the 3'-end of an RNA primer. *Nucleic Acids Res* 1981;9:2433–2453. [PubMed: 6269049]
27. Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke BJ. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* 2006;12:1494–1504. [PubMed: 16790842]
28. Stagno J, Aphasizheva I, Rosengarth A, Luecke H, Aphasizhev R. UTP-bound and Apo structures of a minimal RNA uridylyltransferase. *J Mol Biol* 2007;366:882–899. [PubMed: 17189640]
29. Deng J, Ernst NL, Turley S, Stuart KD, Hol WG. Structural basis for UTP specificity of RNA editing TUTases from *Trypanosoma brucei*. *EMBO J* 2005;24:4007–4017. [PubMed: 16281058]
30. Wang SW, Toda T, MacCallum R, Harris AL, Norbury C. Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. *Mol Cell Biol* 2000;20:3234–3244. [PubMed: 10757807]
31. Stagno J, Aphasizheva I, Aphasizhev R, Luecke H. Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. *Proc Natl Acad Sci U S A* 2007;104:14634–14639. [PubMed: 17785418]
32. Augustin MA, Reichert AS, Betat H, Huber R, Mörl M, Steegborn C. Crystal structure of the human CCA-adding enzyme: insights into template-independent polymerization. *J Mol Biol* 2003;328:985–994. [PubMed: 12729736]
33. Li F, Xiong Y, Wang J, Cho HD, Tomita K, Weiner AM, Steitz TA. Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell* 2002;111:815–824. [PubMed: 12526808]
34. Xiong Y, Li F, Wang J, Weiner AM, Steitz TA. Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol Cell* 2003;12:1165–1172. [PubMed: 14636575]
35. Xiong Y, Steitz TA. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* 2004;430:640–645. [PubMed: 15295590]
36. Tomita K, Fukai S, Ishitani R, Ueda T, Takeuchi N, Vassilyev DG, Nureki O. Structural basis for template-independent RNA polymerization. *Nature* 2004;430:700–704. [PubMed: 15295603]

37. Balbo PB, Bohm A. Mechanism of Poly(A) Polymerase: Structure of the Enzyme-MgATP-RNA Ternary Complex and Kinetic Analysis. *Structure* 2007;15:1117–1131. [PubMed: 17850751]
38. Cho HD, Verlinde CL, Weiner AM. Archaeal CCA-adding enzymes: central role of a highly conserved beta-turn motif in RNA polymerization without translocation. *J Biol Chem* 2005;280:9555–9566. [PubMed: 15590678]
39. Martin G, Keller W. Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. *RNA* 2004;10:899–906. [PubMed: 15146073]
40. Bard J, Zhelkovsky AM, Helmling S, Earnest TN, Moore CL, Bohm A. Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science* 2000;289:1346–1349. [PubMed: 10958780]
41. Martin G, Keller W, Doublié S. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J* 2000;19:4193–4203. [PubMed: 10944102]
42. Martin G, Möglich A, Keller W, Doublié S. Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. *J Mol Biol* 2004;341:911–925. [PubMed: 15328606]
43. Moure CM, Bowman BR, Gershon PD, Quijcho FA. Crystal structures of the vaccinia virus polyadenylate polymerase heterodimer: insights into ATP selectivity and processivity. *Mol Cell* 2006;22:339–349. [PubMed: 16678106]
44. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol* 1984;44:97–179. [PubMed: 6385134]
45. Balbo PB, Toth J, Bohm A. X-ray crystallographic and steady state fluorescence characterization of the protein dynamics of yeast polyadenylate polymerase. *J Mol Biol* 2007;366:1401–1415. [PubMed: 17223131]
46. Batra VK, Beard WA, Shock DD, Krahn JM, Pedersen LC, Wilson SH. Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* 2006;14:757–766. [PubMed: 16615916]
47. Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 1997;36:11205–11215. [PubMed: 9287163]
48. Moon AF, Garcia-Diaz M, Batra VK, Beard WA, Bebenek K, Kunkel TA, Wilson SH, Pedersen LC. The X family portrait: Structural insights into biological functions of X family polymerases. *DNA Repair* 2007;1–17.
49. Johnson L, Liu S, Gershon PD. Molecular flexibility and discontinuous translocation of a non-templated polymerase. *J Mol Biol* 2004;337:843–856. [PubMed: 15033355]
50. Yin YW, Steitz TA. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* 2004;116:393–404. [PubMed: 15016374]
51. Steitz TA. A mechanism for all polymerases. *Nature* 1998;391:231–232. [PubMed: 9440683]
52. Balbo PB, Meinke G, Bohm A. Kinetic studies of yeast polyA polymerase indicate an induced fit mechanism for nucleotide specificity. *Biochemistry* 2005;44:7777–7786. [PubMed: 15909992]
53. Kühn U, Wahle E. Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 2004;1678:67–84. [PubMed: 15157733]
54. Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 2003;4:223. [PubMed: 12844354]
55. Justesen J, Hartmann R, Kjeldgaard NO. Gene structure and function of the 2'-5'-oligoadenylate synthetase family. *Cell Mol Life Sci* 2000;57:1593–1612. [PubMed: 11092454]
56. Smith CA, Baker EN. Aminoglycoside antibiotic resistance by enzymatic deactivation. *Curr Drug Targets Infect Disord* 2002;2:143–160. [PubMed: 12462145]
57. Reuther J, Wohlleben W. Nitrogen metabolism in *Streptomyces coelicolor*: transcriptional and post-translational regulation. *J Mol Microbiol Biotechnol* 2007;12:139–146. [PubMed: 17183221]
58. Kerwitz Y, Kuhn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J* 2003;22:3705–3714. [PubMed: 12853485]
59. Zhelkovsky AM, Kessler MM, Moore CL. Structure-function relationships in the *Saccharomyces cerevisiae* poly(A) polymerase. Identification of a novel RNA binding site and a domain that interacts with specificity factor(s). *J Biol Chem* 1995;270:26715–26720. [PubMed: 7592899]

60. Zhelkovsky A, Helmling S, Moore C. Processivity of the *Saccharomyces cerevisiae* poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol Cell Biol* 1998;18:5942–5951. [PubMed: 9742111]
61. Shi PY, Maizels N, Weiner AM. CCA addition by tRNA nucleotidyltransferase: polymerization without translocation? *EMBO J* 1998;17:3197–3206. [PubMed: 9606201]
62. Betat H, Rammelt C, Martin G, Mörl M. Exchange of regions between bacterial poly(A) polymerase and the CCA-adding enzyme generates altered specificities. *Mol Cell* 2004;15:389–398. [PubMed: 15304219]
63. Wahle E, Rügsegger U. 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev* 1999;23:277–295. [PubMed: 10371034]
64. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J* 2004;23:616–626. [PubMed: 14749727]
65. Wahle E. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 1991;66:759–768. [PubMed: 1878970]
66. Helmling S, Zhelkovsky A, Moore CL. Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol Cell Biol* 2001;21:2026–2037. [PubMed: 11238938]
67. Suh N, Jedamzik B, Eckmann CR, Wickens M, Kimble J. The GLD-2 poly(A) polymerase activates gld-1 mRNA in the *Caenorhabditis elegans* germ line. *Proc Natl Acad Sci U S A* 2006;103:15108–15112. [PubMed: 17012378]
68. Liu Q, Greimann JC, Lima CD. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* 2006;127:1223–1237. [PubMed: 17174896]
69. Vanacova S, Stefl R. The exosome and RNA quality control in the nucleus. *EMBO Rep* 2007;8:651–657. [PubMed: 17603538]

A. MNT with ATP modeled

B. Bovine PAP- α C. *T. brucei* RET2 D. *A. fulgidus* CCAtr with tRNA **Fig. 1.**

Structures of selected RNA dependent nucleotidyl transferases. (A) MNT (light gray; PDB accession 1NO5; [7]) with ATP (green) and Mg^{2+} (gray sphere) modeled from bovine PAP structure (PDB accession 1Q78). Amino acid ligands relevant for ATP binding are depicted with nitrogen in blue and oxygen in red. (B) Bovine PAP- α structure (PDB accession 1Q78; [42]) with bound ATP (red) and Mg^{2+} (gray). Color code for the domains are: orange for catalytic domain, dark blue for central domain and green for RNA binding domain. (C) Editing TUTase RET2 from *T. brucei* (PDB accession 2B56; [29]). Domains functionally correspond to catalytic domain (orange), central domain (dark blue) and RBD (green) of bovine PAP. (D) Class I CCAtr of *Archaeoglobus fulgidus* with bound tRNA (gray; PDB accession 1SZ1; [35]). The colors correspond to the head (catalytic) domain (orange), neck domain (dark blue), body (green) and tail domain (red).

Fig. 2.

Simplified phylogenetic tree of NTrs relevant to this review. Class I rNTrs include the canonical eukaryotic PAPs (euk PAP) and non-canonical PAPs (TRF) and TUTases (TUT), which are related to the archaeal CCATrs (arch CCA). Class II rNTrs include bacterial CCATrs (bact CCA) and PAPs (bact PAP), eukaryotic CCATrs (euk CCA), plant PAPs and CCATrs (plant PAP CCA), and A-adding (A-add) and CC-adding (CC-add) enzymes found in some bacteria. Minimal nucleotidyl transferases (MNT) are the ancestors of the above rNTrs.

Fig. 3.

Active site structures of five representative rNTrs. Amino acid ligands are shown in yellow color with nitrogen in blue and oxygen in red, except in E, where amino acids are shown in orange if belonging to the catalytic domain and blue when from the central domain. Metal ions are depicted as gray spheres and waters as blue spheres. Nucleotides are colored green for ATP, pink for CTP and cyan for UTP or UMP. Only relevant waters are depicted. Hydrogen bonds are indicated as green dotted lines and metal coordination as red dotted lines. π -stacking interactions are visualized by red double bars and stacking interactions between adenine or uracil bases and amino acids as red double arcs in panels B, D, E and F. Incoming nucleotides are labeled ATP or UTP. (A) Minimal RNA uridylyl transferase TUT4 from *T. brucei* (PDB

accession 2IKF; [28]) with bound UTP and Mg^{2+} . (B) TUT4 from *T. brucei* (PDB accession 2Q0F-A; [31]) with bound UTP, UMP (primer mimic) and two Mg^{2+} ions. (C) CCA-adding enzyme of *A. fulgidus* with bound ATP, Mn^{2+} and primer (PDB accession 1TFW; [35]). Only four nucleotides at the 3' end of the primer tRNA are shown and labeled as C72, A73 (D; discriminator), C74 and C75. (D) *Vaccinia virus* PAP in complex with ATP (VP55; PDB accession 2GA9; [43]). (E) Bovine PAP- α active site with bound ATP and Mg^{2+} (PDB accession 1Q78; [42]). (F) *S. cerevisiae* PAP D1154A mutant bound to ATP and (A)₅ primer; only residues A-1 to A-4 of the primer are shown. Hydrogen bonds to 2' hydroxyl groups of primer residues A-1 to A-3 are indicated (PDB accession 2Q66; [37]).

Fig. 4.

Active site geometry of Pol β (PDB accession 2FMS; [46]). Metal coordinating amino acids are shown in yellow color with nitrogen in blue and oxygen in red, metal ions as gray spheres and waters as dark blue spheres. Only residues 9 and 10 of the primer are shown (P9 and P10). In this ternary complex structure the incoming nucleotide is non-hydrolyzable, which allowed the use of a natural deoxynucleoside monophosphate at the 3' end of the primer (dTMP). The 3' OH participates in the ligation of the metal ion. The template strand is not shown for clarity.

Fig. 5.

Examples of predicted complexes of rNTrs with specificity or processivity factors and RNA substrates (black lines). (A) Vaccinia virus PAP catalytic subunit (VP55) and processivity factor VP39 in complex with RNA primer [43]. (B) Mammalian canonical PAP consisting of catalytic domain (PAP-cat) with hidden central domain and RNA binding domain (RBD) in complex with FIP1 (green) and PABPN1 (yellow) and with bound RNA primer [53,58]. (C) Cytoplasmic PAP GLD-2 of *C. elegans* in complex with the RNA binding subunit GLD-3 and RNA [16]. (D) Trf4 and the RNA binding specificity factor Air1 or Air2 of *S. cerevisiae* and tRNA [19].

Table 1

Properties of some rNTrs of known structure

Name	PDB	Processivity	Additions	Flexibility	Induced fit	Specificity
Yeast PAP	2Q66	+	~50	+	+	+
Bovine PAP	1Q78	+	~250	+	+	-
Vaccinia PAP	2GA9	(+/-)	≤30	-	-	+
Class I CCAtr	1TFW-B	(+)	≤3	+	+	-
Class II CCAtr	1MIW	(+)	≤3	(-)	(-)	+
RET2	2B56	-	1	+	(+)	+
7bTUT4	2Q0F-A	+	>200	+	(+)	+

Data were collected from references to published structures and are in parentheses when estimated. The PDB column contains one selected most informative Protein Data Bank (PDB) file, although in most cases additional files are available. Processivity indicates multiple nucleotide additions without dissociation of the primer *in vivo* or *in vitro* with processivity factors present. Flexibility refers to observed or measured movement of the protein domains, often in connection with an induced fit mechanism (see text). Addition indicates the number of nucleotides added per RNA substrate by the rNTrs *in vivo*. Induced fit refers to an induced fit mechanism observed *in vitro*. Specificity refers to apparent substrate specificity seen in structures of binary or ternary complexes listed in the PDB column.