Multiple personalities of the RNA polymerase active centre

Nikolay Zenkin

Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX, UK

Transcription in all living organisms is accomplished by highly conserved multi-subunit RNA polymerases (RNAPs). Our understanding of the functioning of the active centre of RNAPs has transformed recently with the finding that a conserved flexible domain near the active centre, the trigger loop (TL), participates directly in the catalysis of RNA synthesis and serves as a major determinant for fidelity of transcription. It also appears that the TL is involved in the unique ability of RNAPs to exchange catalytic activities of the active centre. In this phenomenon the TL is replaced by a transcription factor which changes the amino acid content and, as a result, the catalytic properties of the active centre. The existence of a number of transcription factors that act through substitution of the TL suggests that the RNAP has several different active centres to choose from in response to external or internal signals.

A video of this Prize Lecture, presented at the Society for General Microbiology Annual Conference 2014, can be viewed via this link: https://www.youtube.com/watch?v=79Z7iXVEPo4

TRIGGER LOOP: A NEW CATALYTIC DOMAIN OF THE RNA POLYMERASE ACTIVE CENTRE

Multi-subunit RNAPs are the enzymes that perform transcription in all living organisms. RNAPs are highly conserved, and emerged before the divergence of the bacteria and archaea/eukaryote lineages. All RNAPs share the invariantly conserved catalytic core of five subunits: β, β', 2σ and ω (bacterial nomenclature is used throughout). The two largest subunits, β and β', form the catalytic cleft where the reaction of addition of nucleoside monophosphates (NMPs) to the growing RNA takes place. As with many other nucleic-acid-managing enzymes (Steitz & Steitz, 1993) and all nucleic-acid-polymerizing enzymes (Steitz, 1998), RNAP uses a two-metal-ion (Mg$^{2+}$) mechanism to catalyse the phosphotransfer reaction. One of the metal ions, Mg$^{2+}$I, is chelated by the invariant triad of aspartates of the β' subunit, whilst the other one, Mg$^{2+}$II, is brought by substrates, e.g. by the incoming nucleoside triphosphate (NTP). Mg$^{2+}$I stabilizes the negative charge on the attacking oxygen of the hydroxyl group of the 3' NMP of RNA. Mg$^{2+}$II assists the leaving of the pyrophosphate, and both metal ions ligate the non-bridging oxygen to stabilize the pentacovalent transition state. Substitutions in the aspartate triad lead to almost full inactivation of the enzyme (Zaychikov et al., 1996).

The emergence of the first crystal structure of multi-subunit RNAP (Zhang et al., 1999) ignited structure-based functional analysis of the enzyme by many researchers. We were interested in the disordered loop (referred to initially as the G-loop and, later, as the TL) of the highly conserved G domain of the β' subunit. In the structure, this disordered region was located >20 Å from Mg$^{2+}$I. However, we found that deletion of this loop had a dramatic effect on catalysis by slowing down NMP addition ~10$^4$ times – an effect close to that of mutations in the aspartate triad. Three years later we found that the antibiotic streptolydigin had a very similar effect to that of the deletion of the TL (Temiakov et al., 2005). Streptolydigin did not inhibit mutant RNAP that lacked the TL (ΔTL RNAP), although it was still able to bind to it, suggesting that streptolydigin imposes inhibition by acting on the TL. The crystal structure of RNAP with streptolydigin bound to it indeed revealed that streptolydigin, whilst binding ~20 Å from Mg$^{2+}$I, fixes the TL in a certain conformation (Temiakov et al., 2005). All these observations led us to the conclusion that the TL must play some critical role during catalysis. Indeed, the active centre of RNAP in complex with streptolydigin appeared to be very ‘open’ with only minimal contacts to the substrates from the surrounding protein (Temiakov et al., 2005) (Fig. 1a). Such conformation is unlikely to be catalytically efficient given the free access for external water molecules that may poison the reaction. A similar observation was made with yeast polymerase II elongation complex where the bound NTP appeared to be bound in a catalytically inactive form, suggesting the requirement for some rearrangement of

Abbreviations: NMP, nucleoside monophosphate; NTP, nucleoside triphosphate; PDB, Protein Data Bank; RNAP, RNA polymerase.
the active centre prior to catalysis (Kettenberger et al., 2004). We therefore hypothesized that a closure of the active centre must take place for catalysis to happen and that the TL must somehow be responsible for the closure. This was the first evidence for a possible role of the TL in catalysis.

The crystal structures of yeast and bacterial RNAPs (Vassylyev et al., 2007; Wang et al., 2006) supported the hypothesis about TL participation in the closing of the active centre and catalysis. Both structures showed the TL in the folded conformation, when several residues of the TL come into touching distance with the reactants in the active site. Later, our and Landick’s groups showed that His1242 and R1239 (Thermus aquaticus numbering) of the TL participated directly in catalysis of the phosphodiester bond formation by stabilizing the transition state of the reaction (Yuzenkova et al., 2010; Zhang et al., 2010). The TL exists in two functional states: catalytically inactive open/unfolded (which is stabilized by streptolydigin) and catalytically active closed/folded. Unfolding of the TL may be required for the NTP substrates to enter and pyrophosphate to leave the active centre through the secondary channel, which, in the folded state of the TL, is closed (Fig. 1, compare a and b). In addition, Kashlev’s and our laboratory showed that the ability of the TL to adopt two conformations is the prerequisite for the high accuracy of RNA synthesis (Kireeva et al., 2008; Yuzenkova et al., 2010). In the ‘induced fit’ mode, the correct NTP induces folding of the TL, which, in turn, provides the physical surface complementary to the NTP and thus stabilizes the

Fig. 1. Multiple active centres of multi-subunit RNAPs. (a–g) Amino acids of β (brown) and β′ (beige) surrounding catalytic Mg²⁺ ions (green) are shown as spheres based on the structure of the bacterial elongation complex [Protein Data Bank (PDB) ID: 2O5J]. (a, b) Incoming NTP (orange) and 3′ NMP of RNA (red) are from PDB ID: 2O5J. RNA nucleotides at +1 and −1 positions (red in c–g) and backtracked 3′ NMP (dark grey spheres in c) are from the structure of the backtracked elongation complex of polymerase II (PDB ID: 3GTJ). DNA and the rest of the RNA are omitted for clarity. The closed TL (grey spheres in b, c) is from PDB ID: 2O5J. The open TL (grey spheres in a, d–g) is from PDB ID: 2BE5. Structures of GreA (PDB ID: 1GRJ), GreB (PDB ID: 2P4V), Gfh1 (PDB ID: 3AOH) and DksA (PDB ID: 1TJL) (grey spheres in d–g) were aligned with the elongation complex according to the structure of the elongation complex with Gfh1 (PDB ID: 3AOH).
transition state of the reaction. However, due to the altered geometry of the non-complementary NTP in the active centre, the TL is not able to fold properly and thus to perform catalysis (Yuzenkova et al., 2010). This increases discrimination against non-complementary NTPs to $10^5$–$10^6$-fold as compared with 10–100-fold for the ‘motionless’ active centre (such as in ΔTL RNAP or RNAP inhibited by streptolydigin) (Yuzenkova et al., 2010). The TL is also solely responsible for the discrimination against dNTPs, apparently being able to fold properly only in the presence of the 2’ OH of the incoming nucleotide (Yuzenkova et al., 2010). Together, the above studies revealed that the TL is a flexible catalytic module of the active centre which is essential, along with catalytic Mg$^{2+}$ ions, for phosphodiester bond synthesis.

**TL ACTIVE CENTRE IS FLEXIBLE FOR TUNING**

In addition to NMP addition (and its direct reversal, i.e. pyrophosphorolysis), the RNAP active centre can hydrolyse the phosphodiester bonds. Hydrolysis is required to excise incorrectly incorporated NMPs and to rescue RNAP from the backtracked state. During backtracking or after misincorporation, the 3’ RNA disengages from the template DNA and active centre, whilst RNAP shifts backwards, thus positioning an upstream RNA phosphodiester bond in the active centre. In such a conformation, NMP addition cannot take place and synthesis is stalled. However, hydrolysis of the phosphodiester bond in such a complex can re-establish the RNA’s 3’ end in the active centre so that it can again be elongated. Hydrolysis is catalysed by the same Mg$^{2+}$ ions of the active centre (Sosunov et al., 2003). Hydrolysis also requires the catalytic domain TL (Yuzenkova & Zenkin, 2010). The folded TL accelerates hydrolysis up to ~500-fold by participating directly in the reaction as a general acid/base (through its His1242). Intriguingly, we showed that the reaction in the 1 bp backtracked state can be assisted by the 3’ NMP of the RNA transcript – the phenomenon we referred to as transcript-assisted hydrolysis (Zenkin et al., 2006). RNA’s 3’ NMP disengages from the DNA template base and is flipped to come closer to the catalytic Mg$^{2+}$ (dark grey in Fig. 1c). From this conformation, the 3’ NMP helps RNAP to chelate catalytic Mg$^{2+}$II, to position the attacking water molecule and possibly participate in hydrolysis of the second phosphodiester bond as a general acid/base (Zenkin et al., 2006). The misincorporated or backtracked 3’ NMP can thus be seen as helping to excise itself from the transcript. The ribozyme-like manner of the transcript-assisted hydrolysis suggests that this may be an ancient mechanism used by the common ancestor of multi-subunit RNAPs. Indeed, we observed recently that the 3’ NMP also participates in intrinsic hydrolysis by eukaryotic polymerase II (Nielsen & Zenkin, 2013).

Cooperation between the catalytic module TL and 3’ NMP can be seen as a composite catalytic module, ‘TL + 3’ NMP’, which participates in the hydrolysis along with the Mg$^{2+}$ ions of the active site. Phosphate, sugar and the base of the 3’ NMP are involved in the reaction differently depending on the identity of the NMP and the corresponding base in the template DNA strand (Zenkin et al., 2006). There are thus a number of different ‘TL + 3’ NMP’ catalytic modules with different catalytic properties.

**REPLACEMENT OF THE TL BY TRANSCRIPTION FACTORS AND SWITCHING OF RNAP CATALYTIC ACTIVITIES**

If not resolved through hydrolysis, backtracked complexes form temporary obstacles for other RNAPs and the replication fork, which may be detrimental to cells (unpublished) (Dutta et al., 2011). In addition, further extension in misincorporated complexes would lead to erroneous transcripts. Intrinsic ‘TL + 3’ NMP’-catalysed hydrolysis is relatively slow and it is no surprise that evolution led to the emergence of specific factors to assist hydrolysis. Cleavage factors bind in the vicinity of the active centre, and help RNAP to chelate Mg$^{2+}$II and, possibly, position the attacking water molecule (Laptenko et al., 2003; Sosunova et al., 2001). We investigated how the bacterial Gre factor cooperates with the ‘TL + 3’ NMP’ catalytic module to enhance the rate of the reaction. Surprisingly, we found that the Gre-assisted reaction did not require the presence of the TL at all, as ΔTL RNAP was as efficient in the Gre-assisted hydrolysis as the WT RNAP (Roghanian et al., 2011). The characteristics of the Gre-assisted hydrolysis were different from those of the reaction catalysed by the TL (irrespective of the identity of the 3’ NMP), and were very similar for ΔTL and WT RNAPs (Roghanian et al., 2011). This observation led us to the proposal that Gre switches off the TL-assisted catalysis and manifests new catalytic properties to the active centre. We proposed that Gre physically replaces the TL in the active site of RNAP. This proposition found support with publication of the crystal structure of the elongation complex with a homologue of Gre, Ghf1, in which the TL is substituted by the long domain of Ghf1 (Tagami et al., 2010). TL in this complex is in its inactive unfolded state. As can be seen from Fig. 1(b–d), Gre that substitutes for the TL must change substantially the amino acid content of the active site as compared with the folded TL. This would mean effectively that the ‘TL (or rather TL + 3’ NMP) active centre’ is exchanged for the ‘Gre active centre’. The Gre catalytic module is at least 100 times more efficient in phosphodiester bond hydrolysis as the TL + 3’ NMP module and is almost as efficient as the TL catalytic module is in NMP addition. Importantly, we showed that substitution of the TL by Gre happens in a controlled manner, i.e. only when RNAP undergoes misincorporation or backtracking (Roghanian et al., 2011). After cleavage takes place, the Gre catalytic module is substituted back by the TL catalytic module, so the synthesis can continue (Roghanian et al., 2011). This may be critical, as high hydrolytic activity of the Gre catalytic module, if not controlled, would lead to...
rapid degradation of the nascent transcript. Interestingly, the switch between Gre and TL may happen without dissociation of Gre from the elongation complex (Yuzenkova et al., 2012).

**MULTIPLE ACTIVE CENTRES OF RNAP**

Gre is not the only catalytic module besides the TL or TL + 3’ NMP modules (Fig. 1). For example, *Escherichia coli* has two Gre factors, GreA and GreB, the properties of which differ somewhat, although both catalyse phosphodiester bond hydrolysis; GreA is more effective on short backtrack complexes, whilst GreB is more proficient on deep backtrack complexes (Borukhov et al., 1993). As seen from Fig. 1(d, e), there is a marked difference in amino acid composition between active centres made up of GreA and GreB. The Gre homologue factor Gfh1 substitutes for the TL as revealed by crystallography (Tagami et al., 2010) (Fig. 1f). The properties of the Gfh1 active centre are not yet clear. It was shown to have a relatively low inhibitory effect on TL-dependent activities (Laptenko et al., 2006). However, by analogy with Gre, the switch between TL and Gfh1 active centres may happen in a controlled manner in response to some signals which have not yet been found. Even more mysterious is the role of the possible active centre formed by the Gre homologue DksA (Fig. 1g). DksA is a global regulator of transcription. The only prominent effect that was observed with DksA was destabilization of promoter complexes (Paul et al., 2004). Recently, DksA was implicated in the control of elongation pausing (Zhang et al., 2014). The molecular mechanisms of the actions of DksA are not known. Although DksA is believed to bind to RNAP in a manner similar to Gre, no strong effects on TL-catalysed reactions or imposition of new catalytic activities has been observed so far. One possibility is that, as in the case of Gfh1, signals that control imposition of the DksA active centre have also not yet been found. The existence of other homologues of Gre, such as Yacl and TraR in *E. coli*, implies the existence of more active centres and activities of RNAPs. Interestingly, small molecules may further modify the mobile active centres. The antibiotic tagetitoxin binds in the vicinity of the catalytic residues of the TL (Artsimovitch et al., 2011; Vassylyev et al., 2005). Whilst not affecting the catalytic properties of the TL active centre (Yuzenkova et al., 2013), tagetitoxin modifies it so that, after phosphodiester bond formation, RNAP with the ‘TL+ tagetitoxin’ active centre has a delayed translocation (Yuzenkova et al., 2013). The antibiotic Microcin J25 binds in the secondary channel and likely modifies the TL active centre so that RNAP becomes more responsive to pauses in transcription (Adelman et al., 2004), although the mechanism is not yet clear. A schematic view of switching between the active centres is shown in Fig. 2.

Unrelated to Gre, but homologous to each other, the eukaryotic andarchaea cleavage factors (TFS in archaea, TFIIIS in RNAP II, C11 in RNAP III and 12.6 in RNAP I) also impose their catalytic activities by substituting for the TL (Cheung & Cramer, 2011; Engel et al., 2013). Therefore, the switching of the TL active centre by external catalytic modules is a common feature of multi-subunit RNAPs from different domains of life, even though the external catalytic modules may be evolutionary and structurally unrelated. The existence of multiple exchangeable active centres (Fig. 2) is a highly unusual property of proteinaceous enzymes. Evolution of such a feature in the case of RNAPs may be explained by the utmost importance of the functions performed by these enzymes – not only highly regulated synthesis of correct mRNAs, ribozymes and structural RNAs, but also possibly replication of the RNA genome of the RNA-protein-based common ancestor of all living organisms.

**ACKNOWLEDGEMENTS**

The work was supported by the UK Biotechnology and Biological Sciences Research Council, the European Research Council and a Wellcome Trust Investigator Award.

**REFERENCES**


