Dissertation
submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by

M.Sc. Mathias Girbig
born in: Malsch, Germany
Oral examination: 26th of October 2021

Structural analysis of human and yeast RNA polymerase III

Referees:
Prof. Dr. Irmgard Sinning
Dr. Martin Beck

Summary

The 17-subunit RNA polymerase III (Pol III) complex transcribes short, structured, and essential RNAs like tRNAs. Mis-regulated Pol III is associated with various genetic diseases including cancer and rare developmental diseases. Pol III is also highly specialized to produce short and abundant transcripts, and a key specialization is transcription termination. Pol III terminates on poly-thymidine (dT) sequences on the non-template (NT) strand of the DNA in the absence of any co-factors, which allows fast recycling of Pol III to start a new transcription cycle. Despite recent advances in the structural characterization of yeast Pol III, little is known about the structure of human Pol III, which hinders studying the role of Pol III in human and health and disease. Furthermore, the molecular mechanism of Pol III transcription termination is not known. Here, I present molecular insights into the structure and function of human and yeast Pol III, which I obtained via cryo-electron microscopy (cryo-EM), to address these questions.

I first report the cryo-EM structures of native human Pol III reaching a resolution up to 2.8 Å. The structures shed light on multiple structural features that are different from yeast Pol III. The human RNA-cleaving subunit RPC10 adopts different conformations and can insert into the Pol III active site. A C-terminal extension of subunit RPC5 that is absent in yeast Pol III binds the Pol III core and might function in transcription initiation of metazoan Pol III. An iron-sulphur cluster bound to the human Pol III heterotrimer subcomplex links the Pol III heterotrimer to the Pol III core. The subunit RPC7α, which is associated to cancer, docks onto the Pol III clamp domain and might, thereby, interfere with the inhibition by the repressor MAF1. I also demonstrate that the human Pol III complex can be produced recombinantly, which will enable an in-depth functional analysis of structure and regulation of human Pol III.

Secondly, I report the cryo-EM structure of the yeast Pol III pre-termination complex (PTC) at 2.8 Å resolution. The structure reveals how Pol III recognizes the poly-dT termination signal. The formation of multiple hydrogen-bonds between a set of highly conserved residues in subunit C128 and the termination signal explains how Pol III pauses prior to termination, which I confirm via structure-function studies. NT-strand recognition induces a contraction of the Pol III core, which is reversed in the absence of the C53-C37 heterodimer, explaining its key function in transcription termination. Finally, the structure of the Pol III PTC in the presence of nucleotides shows how Pol III gets trapped in a non-productive state once it reaches the termination signal. Based on these findings, I propose a structure-based mechanism of Pol III transcription termination.

Zusammenfassung

Der RNA Polymerase III (Pol III) Komplex, bestehend aus 17 Untereinheiten, transkribiert kurze, strukturierte und essentielle RNAs, wie zum Beispiel tRNAs. Falsch regulierte Pol III ist mit verschiedenen genetischen Krankheiten, einschließlich Krebs und seltene Entwicklungsstörungen, assoziiert. Pol III ist hochspezialisiert auf die Produktion von kurzen und reichlich vorhandenen Transkripten und eine Schlüsselspezialisierung ist die Terminierung der Transkription. Pol III terminiert an poly-Thymidin (dT) Sequenzen auf dem nicht-codogenem Strang der DNA in der Abwesenheit jeglicher Cofaktoren, was eine schnelle Wiederverwendung der Pol III ermöglicht, um einen neuen Transkriptionszyklus zu starten. Trotz der jüngsten Fortschritte bei der strukturellen Charakterisierung der Hefe Pol III, ist wenig über die Struktur der humanen Pol III bekannt, was es erschwert, die Rolle der Pol III in der menschlichen Gesundheit und Krankheit zu untersuchen. Des Weiteren ist der molekulare Mechanismus der Pol III Trankriptionsterminierung nicht bekannt. Hier präsentiere ich molekulare Einblicke in die Struktur und Funktion der Pol III von Mensch und Hefe, welche ich mithilfe der Kryo-Elektronenmikroskopie (Kryo-EM) erzielt habe, um diese Fragen zu addressieren.

Zuerst stelle ich die Kryo-EM Strukturen der nativen humanen Pol III vor, die eine Auflösung bis zu 2.8 Å erreichen. Die Strukturen geben Aufschluss auf mehrere Elemente, die sich von der Hefe Pol III unterscheiden. Die RNA-schneidende Untereinheit RPC10 nimmt verschiedene Konformationen an und kann in das aktive Zentrum der Pol III hineinreichen. Eine C-terminale Verlängerung von RPC5, welche in der Hefe nicht vorhanden ist, bindet den Kern der Pol III und könnte eine Funktion bei der Transkriptionsinitiation von Tieren spielen. Ein Eisen-Schwefel Cluster das an das humane Heterotrimer bindet, verknüpft das Pol III Heterotrimer mit dem Pol III Kern. Das Krebs-assoziierte RPC7α dockt an die Pol III *Clamp*-Domäne und könnte dadurch die Inhibition durch den Repressor MAF1 behindern. Ich demonstriere auch, dass der Pol III Komplex rekombinant hergestellt werden kann, was eine detaillierte funktionale Untersuchung von der Struktur und Regulierung der humanen Pol III ermöglichen wird.

Als zweites präsentiere ich die Kryo-EM Struktur des Pol III Prä-Terminations-Komplexes (PTC) bei einer Auflösung von 2.8 Å. Die Struktur zeigt, wie Pol III das poly-dT Terminationssignal erkennt. Das Ausbilden mehrerer Wasserstoffbrückenbindungen zwischen stark konservierten Resten der Untereinheit C128 und dem Terminationssignal erklärt, wie Pol III vor der Termination pausiert, was ich mit Struktur-Funktionsuntersuchungen untermauere. Die Erkennung des nicht-codogenen Stranges löst ein Zusammenziehen des Pol III Kerns aus, welches durch die Abwesenheit des C53-C37 Heterodimers reversiert wird. Dies erklärt die Funktion des Heterodimers in der Transkriptionsterminierung. Abschließend zeigt die Struktur des Pol III PTC in Anwesenheit von Nukleotiden, wie Pol III in einem unproduktivem Zustand gefangen wird, sobald sie das Terminationssignal erreicht. Basierend auf diesen Erkenntnissen schlage ich einen strukturbasierten Mechanismus zur Pol III Transkriptionsterminierung vor.

Peer-reviewed articles

Girbig M*, Misiaszek AD*, Vorländer MK, Lafita A, Grötsch H, Baudin F, Bateman A, Müller CW (2021). Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. Nat Struct Mol Biol; 28(2): 210-19

* Equal contribution

Preprints

Misiaszek AD, **Girbig M**, Grötsch H, Baudin F, Lafita A, Murciano B, Müller CW (2021). Cryo-EM structures of human RNA polymerase I. bioRxiv (in revision for Nat Struct Mol Biol)

Manuscripts in preparation

Girbig M, Grötsch H, Xie J, Baudin F, Libri D, Porrua O, Müller CW. Structural basis of RNA polymerase III transcription termination.

Acknowledgements

First of all, I would like to express my deepest gratitude to Dr. Christoph Müller for giving me the chance to work in his group and on these exciting projects. Christoph has been an amazing advisor and was incredibly supportive over the entire course of my PhD. I highly appreciate and respect his way of supervision that included giving me a lot of freedom to explore my own ideas while still ensuring that I'm on the right track. I will always be proud to be an alumnus of the Müller lab once I'm leaving EMBL to continue my academic journey.

I would like to thank all the past and present members of the Müller group for a lot of assistance and for making the time in the lab so enjoyable. I'm also particularly thankful to four members of the Müller lab who are listed in an alphabetical order of their surnames because it would be impossible for me to choose whom I should thank the most: Dr. Florence Baudin, Dr. Helga Grötsch, Agata Misiaszek, and Dr. Matthias Vorländer.

Florence taught me a lot about nucleic acid biochemistry, and the knowledge and skills that I acquired from her were particularly important in the beginning of my PhD. Florence also performed some key experiments, and it has always been great to discuss and evaluate these experiments together on such an intellectual but also humorous level.

Helga, our lab manager, was incredibly helpful, and I will remember her as the lab's knight in the shining armour as I'm convinced many experiments would have failed without her crucial expertise in molecular biology. Helga's dedication to assist with the handling of human and yeast cells and their genomic engineering was critical for this work.

Agata is not only a great colleague but also someone with whom I heavily collaborated with over the course of the last two years. I'm very grateful and impressed that Agata was brave enough to take on the genomic engineering of the HEK293 cells, which was crucial for the structural analysis of human Pol III.

Matthias was very helpful with the cryo-EM sample preparation, data collection, and processing. He also taught me how to build and refine the structural models of macromolecular complexes, which was not only key for this work but will also be extremely valuable for my future research activities.

I would like to thank Wim Hagen and Dr. Felix Weis from the EMBL cryo-EM service platform for their immense support regarding the preparation, data collection, and processing of the cryo-EM data. Likewise, I want like to acknowledge Dr. Thomas

Hoffmann for the valuable IT support he provided. I am also grateful to my friend and former colleague, Dr. Aleix Lafita, who helped with the phylogenomic analysis of Pol III. It has also been a great pleasure to discuss my research with- and receive valuable feedback from Dr. Sebastian Eustermann.

I also want to thank the members of my TAC, Prof. Dr. Irmi Sinning, Dr. Martin Beck, and Dr. Eileen Furlong, for their valuable advice over the course of my PhD project.

I want to thank the Boehringer Ingelheim Fonds, not only for their financial support but also for the fact that I got to know so many truly inspiring fellow colleagues and friends.

I am truly convinced that I would have not entered the EMBL International PhD Programme without the supervision and guidance of the following fantastic mentors who supervised and supported me over the course of the last nine years: Dr. Lisa Henning, Dr. Gesa Albert, Prof. Dr. Christian Freund, Dr. Karine dos Santos, Dr. Matthias Theuser, Prof. Dr. Markus Wahl, Dr. Ursula Egner, Dr. Stephan Grunwald, Dr. Oliver Daumke, Prof. Dr. Ralf Jauch, Dr. Andrew Carter, and Dr. Christoph Müller. Some of them have supervised me for several years and will always be my scientific role models.

Besides, I am also grateful for all the great friends that I got to know throughout my academic journey like those with whom I: can always discuss my research and exchange ideas (Vladimir Ugorets, Leo Kiss, Dr. Tobias Wauer, Prof. Dr. Andreas Boland), lived together (Vladislav Kim), or practiced meditation at EMBL (Can Sonmezer and Jakob Wirbel). The same applies to my two best friends from school, Manuel Buckenmaier and Ann-Kathrin Sobotta, with whom I had far too little contact during my PhD.

Finally, I want to say a big thank you to my family and to my closest relatives: my sister, Dr. Dorothee Childs and her husband Dr. Liam Childs for being great role models and demonstrating that it is possible to be amazing parents and great scientists at the same time; my fiancé, Qitong Zhou, for accepting my dedication to science and for her unconditional love, trust, and support; and of course my parents, Ludwig and Helga Girbig, for their interest in my research activities and their everlasting support, without which this work would not exist.

Contents

Part I.
Introduction

1. Introduction

RNA transcription describes the process of RNA synthesis that is carried out by RNA polymerases and is an universal process in bacteria, archaea, eukaryota, and viruses [1, 2]. Whereas prokaryotes and eukaryotes use DNA-dependent RNA polymerases (RNAPs) to synthesize most[1] of their RNA repertoire, viruses can transcribe RNA by either DNA- or RNA-dependent RNA polymerases depending on whether their genetic material consists of DNA or RNA [2]. Bacterial, archaeal and eukaryotic RNAPs are multisubunit protein complexes [1], except for the nuclear-encoded RNA polymerases in mitochondria and chloroplasts, which are bacteriophage-type single subunit RNA polymerases [4]. This work focuses on the structural analysis of the RNA polymerase III, which is a eukaryotic RNAP that, in some parts, not only shares homology with the other eukaryotic RNAPs but also with those found in bacteria and archaea.

In the first part of the introduction, I will give a general overview about RNA transcription by multi-subunit RNAPs. The general transcription cycle will be briefly introduced, and the molecular basis of the catalytic reaction, upon which the RNA is synthesized will be described. Because multi-subunit RNAPs can be found in all three domains of life, I will outline the commonalities and differences between these RNAPs. The second part of the introduction focuses on Pol III and (i) its function and target genes; (ii) the composition and structure of Pol III itself and its cofactors; (iii) on human Pol III, its conservation to yeast Pol III and its role in health and disease; and (iv) on Pol III transcription termination. The last part of the introduction briefly describes the main technique that I've used for the structural analysis of Pol III: cryo-electron microscopy (cryo-EM). Finally, I will outline the scope of my thesis.

[1] Eukaryotes also contain RNA-dependent RNA polymerases [3]

1.1. RNA transcription by DNA-dependent RNA polymerases

1.1.1. The RNA transcription cycle

DNA-dependent RNAPs use DNA as a template to transcribe the RNA. In all three domains of life, the process of RNA transcription can be divided into three major phases: (i) transcription initiation; (ii) transcription elongation; and (iii) transcription termination (Fig. 1.1 a). The characteristic features of the three phases (reviewed in [5]) are described below:

(i) **Transcription initiation** Transcription begins with the initiation phase, which not only requires the RNAP itself but also additional co-factors such as general transcription factors (TFs) (Fig. 1.1 b). While the number of required TFs and additional co-factors, as well as the extend of structural modulation of the RNAP substantially varies across the different RNAPs, the following steps are applicable to all RNAPs. First, TFs bind the promoter DNA, which often occurs in a step-wise manner. Subsequently, the DNA-unbound (apo) RNAP is recruited to the promoter DNA with the help of the TFs to form the pre-initiation complex (PIC). In the first place, the double-stranded (ds) promoter DNA is closed, which is why this conformation is called closed complex (CC). Next, the promoter DNA is melted into the template (T)- and non-template (NT) strands and the open complex (OC) forms, which typically requires the combined action of the RNAP and additional TFs. In the OC, the two DNA strands are, essentially, pulled into the core of the RNAP and the transcription bubble forms, of which the T-strand engages closely with the active site of the RNAP. Next, the initially transcribing complex (ITC) forms. The RNAP starts to anneal nucleotides, derived from incoming nucleoside triphosphates (NTPs), to the template strand and, thereby, synthesizes an RNA that is complementary to the template-strand. The initial length of the so-formed DNA:RNA hybrid is ca. 6 nucleotides and, once the RNA exceeds a length of typically 8-14 nucleotides, the RNAP escapes the promoter DNA.

(ii) **Transcription elongation** Once the RNAP has escaped the promoter DNA, the elongation complex (EC) forms that translocates along the DNA and actively synthesizes the nascent RNA (Fig. 1.1 c). In the EC, the DNA:RNA hybrid is 8 to 9 base-pairs long, tightly engages with the RNAP active site and is, therefore, usually well resolved in the available EC structures (both obtained via X-ray crystallography and cryo-EM). In contrast, the NT-strand is more flexible and typically not visible in the RNAP EC structures. Due to the translocating motion of the RNAP, the RNA gets extended, and the 5'-end of the nascent transcript protrudes out of the RNA exist channel of the RNAP. During transcription elongation, the RNAP EC

can associate with various elongation factors that ensure that the RNAP remains in a processive RNA-synthesizing state. Still, RNAP may pause during transcription elongation [6], which could be a consequence of a nucleosome (in eukaryotes), the formation of R-loops (DNA:RNA hybrids that form between the nascent RNA and the DNA outside of the transcription bubble), or the misincorporation of the wrong nucleotide into the nascent RNA [7]. The latter event induces backtracking of the RNAP, upon which the 3'-end of the RNA is threaded out of another channel - the funnel - and needs to be cleaved off by RNA-cleavage factors. This process reactivates the RNAP so that it can actively translocate along the DNA towards the end of the transcribed gene.

(iii) **Transcription termination** The transcription cycle ends with termination once the RNAP reaches a termination site (Fig. 1.1 d). With the help of additional termination and pausing factors, the RNAP first pauses and subsequently terminates transcription, upon which the transcription bubble collapses and the nascent RNA and the DNA are released. The terminated apo RNAP can then be recycled to start a new round of transcription. Notably, the number of required pausing- and termination substantially differs across the different RNAP systems. A more detailed description of transcription termination of Pol III is given below.

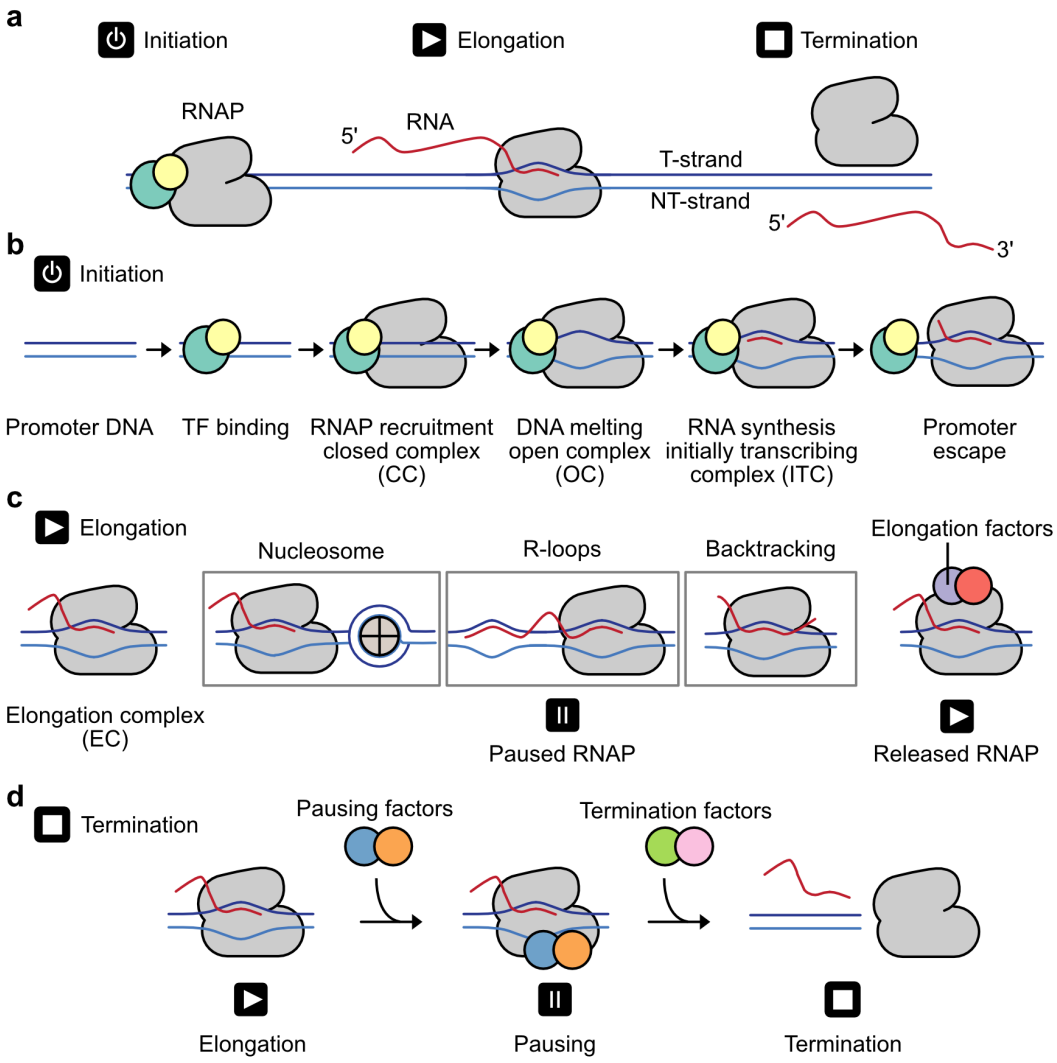

Fig. 1.1. The RNA transcription cycle and its three phases. (a) Schematic of RNA transcription by an RNAP illustrating the major three phases: (i) transcription initiation; (ii) transcription elongation; and (iii) transcription termination. **(b)** Schematic of the transcription initiation phase and its sub-phases. **(c)** Schematic of the transcription elongation phase, during which the RNAP may processively transcribe the RNA in form of the elongation complex (EC) or pause due to various reasons. Three potential elements that could induce pausing are shown: RNAP reaching a nucleosome (in eukaryotes); R-loop formation; Backtracking because of wrong nucleotide incorporation. The paused state may be released by the help of elongation factors that bind the RNAP to restore a processive state. **(d)** Schematic of the transcription termination phase at the termination site. The RNAP first needs to be paused, typically by the help of pausing factors. Subsequently, termination is triggered, typically by the help of termination factors, which destabilize the transcription complex so that the RNAP releases the DNA and the nascent RNA transcript.

1.1.2. Multi-subunit RNAPs in the three domains of life

Whereas prokaryotes (bacteria and archaea) use a single RNAP to synthesize their RNAs, eukaryotes evolved up five specialized RNAPs to split up the task of RNA synthesis [1]. RNA polymerase I (Pol I) transcribes ribosomal RNA (rRNA), RNA polymerase II (Pol II) synthesizes messenger RNAs (mRNAs) and many regulatory RNAs, and RNA polymerase III (Pol III) produces short RNAs such as transfer RNAs (tRNAs). Plants contain two additional RNAPs: RNA polymerase IV (Pol IV) and RNA polymerase V (Pol V). Both enzymes transcribe noncoding RNAs that function in gene silencing [8]. Because these two RNAPs are specific to plants, I will, hereinafter, only refer to those RNAPs that are universal to all eukaryotes when describing the eukaryotic RNAPs: Pol I, Pol II, and Pol III.

A common feature of all multi-subunit RNAPs is the conserved RNAP core that resembles a horseshoe-like shape and harbours the active site in its centre. Bacterial RNAP consists of five subunits (Fig. 1.2 a): β', β, α_1, α_2, and ω. The two β subunits are the largest proteins in the bacterial RNAP and form the active site where the catalytic nucleotide addition reaction occurs that gives rise to the nascent RNA chain (details below). The archaeal RNAP contains up to 12 subunits, listed in Fig. 1.2 b. Whereas 5 subunits of its core are homologous to those found in bacteria, archaea have evolved 7 more subunits, of which Rpo4 and Rpo7 form the stalk domain that protrudes out of the core and is absent in bacteria.

Eukaryotic RNAPs contain between 12- to 17 subunits. They all have a 10-subunit core and 2-subunit stalk. Of these 12 core- and stalk subunits, 11 have orthologs in archaea. Each of these 12 subunits either has a paralogous subunit (e.g. A190, RPB1, C160) in one of the other two eukaryotic RNAPs or is the same in three RNAP structures (e.g. ABC27). The smallest eukaryotic RNAP is Pol II that is made out of 12 subunits and only consists of the stalk and the core, of which 5 subunits are the same between Pol I, Pol II, and Pol III. (Fig. 1.2 c). The second-largest RNAP is Pol I that is build out of 14 subunits (Fig. 1.2 d). Besides the core and stalk, Pol I contains the 2-subunit heterodimer - A49, A34.5 - that is homologous to TFIIF, a general TF of Pol II. Within the core, Pol I not only shares the 5 subunits with Pol I and Pol III but also two additional subunits with Pol III (AC40, AC19). With 17 subunits, Pol III is the largest eukaryotic RNAP (Fig. 1.2 e). Besides the core, stalk and a TFIIF-like heterodimer (C53, C37), Pol III also features the heterotrimer (C82, C34, C31), which shares homology to TFIIE and TFIIF. Both the Pol I and Pol III cores also contain a small subunit (A12.2 in Pol I, C11 in Pol III), of which the N-terminus is homologous to subunit RPB9 in Pol II and the C-terminus is homologous to the Pol II elongation factor TFIIS, which can cleave RNA.

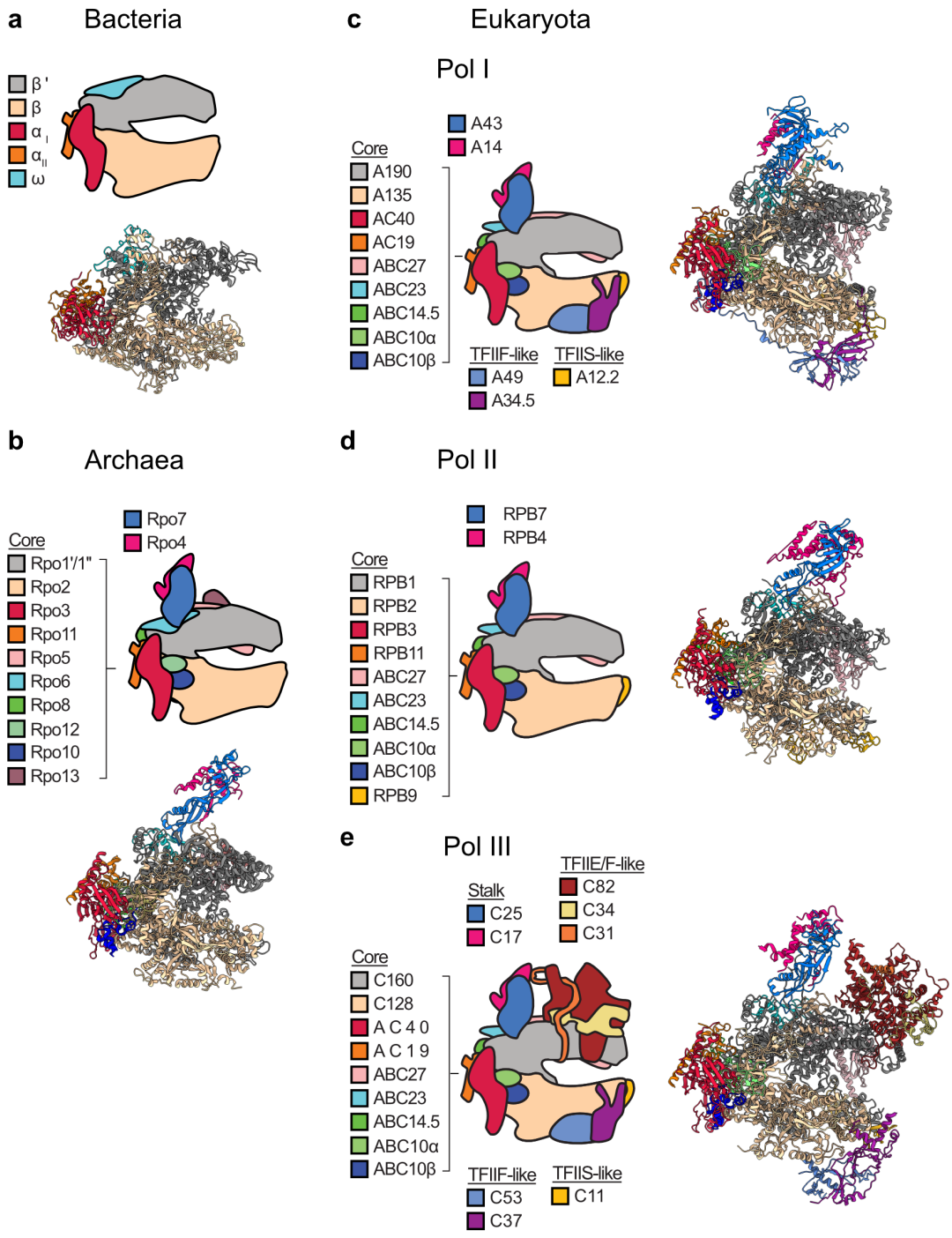

Fig. 1.2. The multisubunit RNAPs in the three domains of life. **(a)-(b)** Schematics (top), structures (bottom), and subunit legends of **(a)** the bacterial RNAP (*Thermus aquaticus*, PDB: 1I6V) and **(b)** the archaeal RNAP (*Sulfolobus shibatae* Protein, PDB: 2WAQ). **(c)-(e)** Schematics (left), structures (right), and subunit legends of **(c)** Pol I (*Saccharomyces cerevisiae*, PDB: 4C3I), **(d)** Pol II (*S. cerevisiae*, PDB: 1WCM), and **(e)** Pol III (*S. cerevisiae*, PDB: 5FJA).

1.1.3. Mechanism of nucleotide addition

The synthesis of RNA requires the addition of a nucleotide to the 3'-end of the RNA chain and is catalysed by the highly conserved active site of the RNAP, which is positioned in the central core of the RNAP (Fig. 1.3 a) (reviewed in [5]). In the active site of the RNAP EC, the dsDNA is melted into NT-strand and T-strand to form a 11-12 nucleotides long transcription bubble (Fig. 1.3 b). The T-strand serves as a template for synthesis of the RNA chain, which is complementary to the T-strand. The DNA:RNA hybrid in the RNAP EC is 8-9 bp long. For the nucleotide addition reaction, an NTP first needs to enter the RNAP core through the funnel domain (see below) and is bound in the RNAP active site at the nucleotide addition site (i+1). The 3'-hydroxyl group of the 3'-end nucleotide (i-1) of the RNA chain get deprotonated and performs a nucleophilic attack on the α-phosphate group of the bound NTP (Fig. 1.3 a). Thereby, a phosphodiester bond between the RNA chain and the bound i+1 nucleotide forms. The former (before the nucleotide reaction) β and γ phosphates of the NTP are released as a pyrophosphate (PP_i). For the nucleotide reaction to take place, two magnesium ions (Mg^{2+}) are required. The metal A magnesium coordinates the 3'-hydroxyl group of the RNA and the α-phosphate group of the bound NTP. The metal B magnesium coordinates the β and γ phosphates of the NTP and is released from the RNAP active together with the PP_i.

The RNA chain elongation requires the translocation of the RNAP along the DNA and is facilitate by the nucleotide-addition cycle and its oscillation between pre- and post-translocation states (Fig. 1.3 c). In the post-translocation state, the RNAP active is not yet loaded with an NTP, and the distance between the 3'-end of the RNA and the transcription fork (the site of the DNA just before it is melted into the two strands) corresponds to 2 nucleotides. The first step is the binding of the NTP to the active, which accommodates the i+1 site. Second, the phosphodiester bond formation takes place, the PP_i is released, and the RNAP EC is in the pre-translocation state. The i+1 site is now occupied by the 3'-end of the RNA, which has been extended by one nucleotide. Third, the RNAP translocates by one base, upon which the i+1 site is cleared and the original 2-nucleotide-long distance between 3'-end RNA and the transcription fork is restored. The translocated state essentially resembles the post-translocated state, and the nucleotide addition cycle can start again. Another step that can take place is backtracking as a result of RNAP pausing or nucleotide-misincorporation. RNAP backtracking implies moving of the RNAP into the opposite direction and, consequently, protrusion of the 3'-end of the RNA out of the active site through the funnel domain. A backtracked state can be released via intrinsic or extrinsic (e.g. via the transcription elongation factor TFIIS in Pol II) RNA cleavage to restore a translocated state.

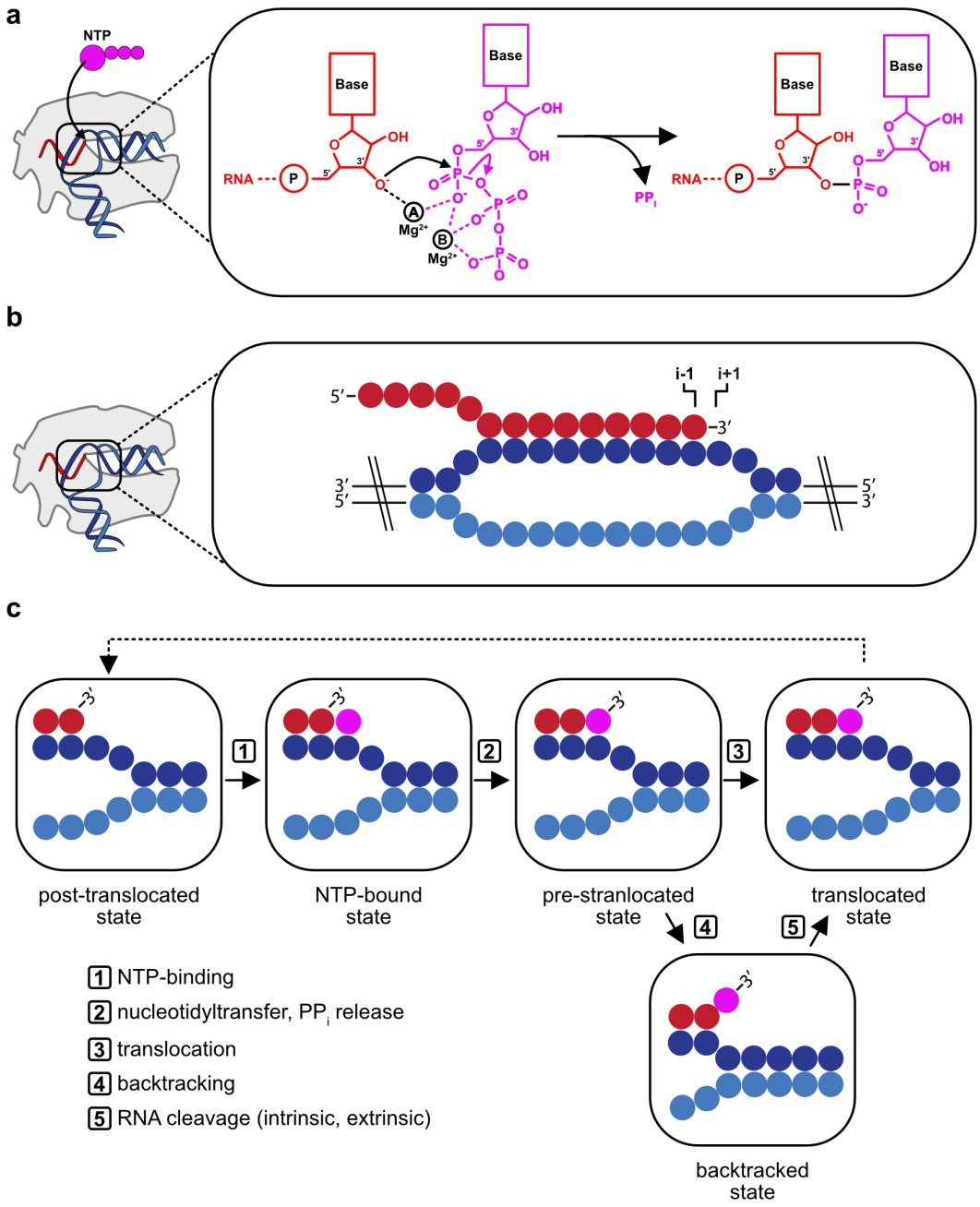

Fig. 1.3. The nucleotide addition cycle. (a) Schematic of the nucleotide addition reaction taking place in the active site of the nucleic acid-bound RNAP (shown on the left). The incoming NTP (pink) is added to the nascent RNA chain (red) via a nucleophilic attack of the 3' hydroxyl-group of the RNA chain on the α-phosphate group of the bound NTP, which results in a newly formed phosphodiester bond and a released diphosphate. Two magnesium-ions (A and B) are required to coordinate the 3'-hydroxyl group and the phosphates of the NTP. **(b)** Schematic of the transcription bubble within the RNAP active site. The nucleotide addition site (i+1) and the site of the 3-base of the nascent RNA (i-1) are labeled. **(c)** Overview of the five steps (1-5) and respective different states during the NTP addition cycle. The incoming NTP is colored in pink. T-strand: template strand; NT-strand: non-template strand.

1.1.4. Structural key elements of eukaryotic RNAPs

RNA transcription is facilitated by the conserved active site of the RNAPs and other structural key elements that can be found in multi-subunit RNAPs. For eukaryotic RNAPs, some of the structural key elements are listed below. Notably, these elements are highly conserved between Pol I, II, and III but are described and displayed in Fig. 1.4 only for Pol II because this was the first eukaryotic RNAP, of which the structure was solved and most of these elements were reported [9–11].

- **DNA-binding cleft**. The DNA-binding cleft is a positively charged cavity that is mostly formed by the largest RNAP subunit RPB1/A190/C160 and binds the incoming downstream dsDNA (Fig. 1.4 a).
- **Clamp domain**. The clamp domain is a mobile element of the RNAP and is mostly formed by RPB1/A190/C160 (Fig. 1.4 b). The clamp can either adopt an open conformation (apo RNAP) or closed conformation (nucleic acid bound RNAP) and serves as binding platform for transcription factors.
- **Stalk**. The stalk is also a mobile element that protrudes out of the RNAP core (Fig. 1.4 b). It is formed by RPB4/7 (Pol II), A14/A43 (Pol I), and C17/C25 (Pol III). The stalk is involved in transcription initiation and serves as an interaction hub for several transcription initiation factors such as TFIIE and the mediator complex in Pol II [12–14], Rrn3 in Pol I [15], or the heterotrimer subcomplex (which can be regarded as an in-built transcription factor) in Pol III [16].
- **Protrusion and lobe**. The protrusion and lobe domains are formed by the second-largest RNAP subunits: Rbp2 (Pol II), A135 (Pol I), and C128 (Pol III) (Fig. 1.4 b). Both domains serve as interaction hubs for transcription factors or RNAP subcomplexes such as TFIIF in Pol II [12, 14], the A49/A34.5 heterodimer in Pol I [17], or the C53-C37 heterodimer in Pol III [16].
- **Wall** The wall is located near the active site, is formed by RPB2/A135/C128, and serves as a polypeptide wall that defines the path of the DNA:RNA hybrid in the active site by forcing the hybrid to bend before exiting the RNAP core (Fig. 1.4 c).
- **Bridge helix**. The bridge helix is formed by RPB1/A190/C160 and is a mobile element that spans the active site of the RNAP (Fig. 1.4 c). It can be partially unfolded (e.g. see [17]) or completely folded (e.g. see [11]) and can adopt a straight or bent conformation, which depends on the RNAP type (e.g. Pol II or Pol I) and conformation (e.g. apo RNAP, EC RNAP). The bridge helix contacts the T-DNA strand and its molecular motion functions in the nucleotide addition cycle and RNAP translocation.

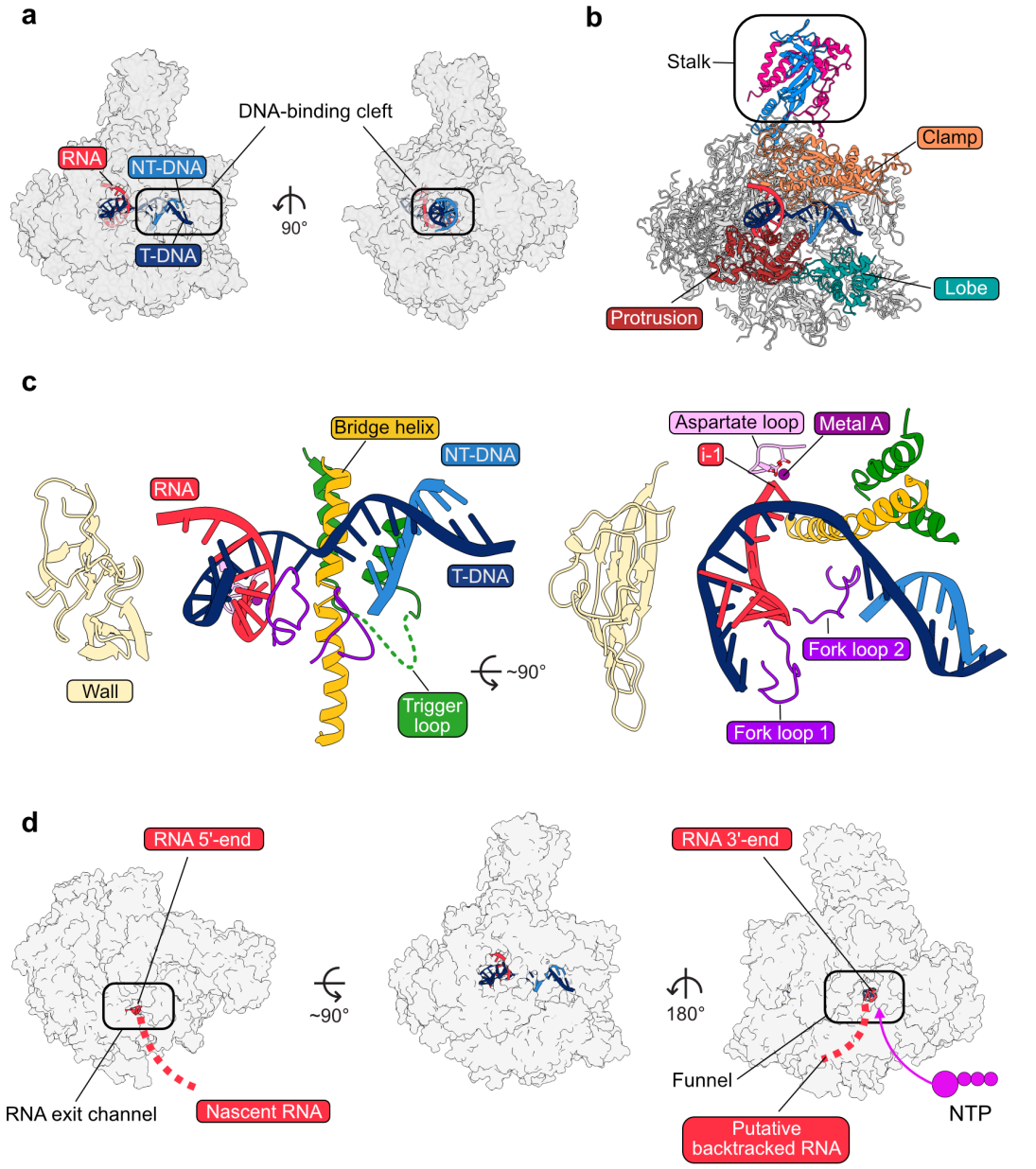

Fig. 1.4. Structural key elements of eukaryotic RNAPs. Shown is the structure of the *S. cerevisiae* Pol II EC (PDB: 1Y1W). **(a)** Front- and side view of the Pol II EC (grey transparent surface) bound to nucleic acids (colored ribbons). The downstream dsDNA is bound by the DNA-binding cleft. **(b)** Side view of the Pol II EC displayed as a ribbon. Characteristic globular key elements are colored and labeled. **(c)** Close-up views onto the Pol II active site with key elements being colored and labeled. **(d)** RNA exit paths of Pol II EC. Three views on Pol II shown as a grey solid surface and bound to nucleic acids shown as colored ribbons. The top (left) view illustrates the RNA exit channel out of which the 5'-end of the nascent RNA chain protrudes. The side view (right) shows the funnel domain that serves as the entry point for the incoming NTPs (magenta schematic representation). In addition, the funnel domain serves as the exit channel for the 3'-end of the backtracked RNA when backtracking occurs.

- **Trigger loop**. The trigger loop, as part of RPB1/A190/C160, is another active site element that participates in the nucleotide addition cycle. It is positioned near the bridge helix and the 3'-end of the RNA chain (Fig. 1.4 c). The trigger loop can adopt multiple conformations (reviewed in [5]) and thereby functions in nucleotide binding and selection and phosphodiester bond formation [18]. Depending on the functional state of the RNAP, the trigger loop either adopts a folded or unfolded conformation.

- **Aspartate loop**. The aspartate loop is positioned in the active centre of the RNAP and is formed by RPB1/A190/C160 (Fig. 1.4 c). It is critical for the nucleotide addition cycle as it contains three aspartate residues that coordinate the metal A (Mg^{2+})-ion, which in turn coordinates the 3'-hydroxyl group of the RNA and the α-phosphate group of the bound NTP [11].

- **Fork loops 1 and 2**. The fork loop 1 and fork loop 2 elements are part of RPB2/A135/C128 (Fig. 1.4 c). Both loops function in stabilizing the transcription bubble as they contact the transcription fork (fork loop 2) and - at least in Pol II - the DNA:RNA hybrid (fork loop 1) [19].

- **RNA exit channel**. The RNA exit channel serves as the exit point for the 5'-end of the nascent RNA chain (Fig. 1.4 d). It is positioned close to the clamp domain and is formed by RPB1/A190/C160 and RPB2/A135/C128.

- **Funnel**. The funnel domain (RPB1/A190/C160) serves as the entry channel for the incoming NTP (Fig. 1.4 d). Furthermore it serves as a secondary RNA exit channel to expel the 3'-end of the backtracked RNA during RNAP backtracking.

1.2. Pol III transcription

1.2.1. Pol III - discovery and target genes

Pol III was identified together with Pol I and Pol II by isolating the three enzymes from nuclear extracts of the sea urchin *Strongylocentrotus purpuratus* that was applied to anion-exchange chromatography [20]. This purification procedure yielded three peaks (I, II, III) that were assigned to Pol I, Pol II, and Pol III, respectively. The three enzymes were later shown - by purification from *Xenopus laevis* [21] and *S. cerevisiae* [22] - to exhibit different molecular weights and numbers of individual subunits. Some subunits were identified to be the same in the three RNAPs whereas others are unique to Pol I, Pol II, and Pol III. Pol III was shown to be the largest RNAP with a total mass of ca. 632 kDa, followed by Pol I with ca. 513 kDa and Pol II with ca. 487 kDa [22]. Shortly after their discovery, the three RNAPs were shown to have distinct functions, and the first described Pol III transcript were tRNAs and the 5S rRNA [23, 24]. The number of reported Pol III transcripts gradually increased over time. The transcripts have in common that they are all relatively short (about 70-369 nucleotides) and exhibit distinct secondary structures. The known Pol III transcripts are listed below:

- **5S ribosomal RNA (5S rRNA)** - The 5S rRNA was - along tRNAs - the first known Pol III transcript [23, 24]. The 5S rRNA is ca. 120 nucleotides long and is a building block of the ribosome. Eukaryotes typically contain several hundreds to thousands of 5S rRNA genes that can be arranged as separate tandem repeats (e.g. in *X. laevis* [25]) or located within a locus that also contains the other 28S, 18S, 5.8S rRNA genes (e.g. in *S. cerevisiae* [26]).

- **transfer RNA (tRNA)** - tRNAs are 70-90 nucleotides long and are loaded with amino acids and, thereby, serve as adaptors during protein synthesis. Together with the 5S rRNA, tRNAs were the first discovered Pol III transcripts [23, 24]. The number of tRNA genes varies across eukaryotes. For example, according to the Genomic tRNA Database [27], *S. cerevisiae*, *Homo sapiens*, and *Arabidopsis thaliana* contains 275, 416, and 580 different tRNA genes, respectively.

- **U6 small nuclear RNA (U6 snRNA)** - The U6 snRNA is part of the U6 small nuclear ribonucleoprotein (snRNP) and, as such, a part of the spliceosome complex, which contains six different snRNAs [28]. Whereas the U1-U5 snRNAs are made by Pol II, the U6 snRNA is transcribed by Pol III [29].

- **Viral associated RNA (VA RNA)** - VA RNAs can be found in many viruses such as adenoviruses and are transcribed at high rates during viral infections by

Pol III [24, 30]. VA RNAs have been shown to be required for the translation of viral mRNAs [31].

- **Epstein–Barr virus-encoded small RNA (EBER)** - EBERs are ca. 170 nucleotides long and are encoded in the genome of the Epstein–Barr virus [32]. EBERs are transcribed by Pol III of the Epstein–Barr virus-infected host cell [33] and are associated with malignant transformation [34]. Their precise role is, however, not yet well understood.
- **Signal recognition particle RNA (7SL RNA)** - The 7SL RNA is a 300 nucleotides long RNA that associates with the signal recognition particle (SRP), which co-translationally targets proteins to the the endoplasmic reticulum [35]. The 7SL RNA has been identified as a Pol III transcript using temperature-sensitive Pol III mutant cells that showed a strong reduction of 7SL RNA levels [36].
- **Ribonuclease P RNA component H1 (H1 RNA)** - With ca. 340 nucleotides, the H1 RNA is the longest Pol III transcript and is the RNA component of the endoribonuclease RNase P, which functions in pre-tRNA processing [37].
- **RNase MRP RNA** - The MRP RNA is a Pol III transcript that is part of the MRP ribonucleoprotein and functions in rRNA processing [38].
- **Vault RNA (vRNA)** - vRNAs are 89-141 nucleotides long and associate with vaults, which are large ribonucleoproteins with a mass of 13 MDa [39]. The precise role of the vRNAs is not well understood but one vRNA transcript (vRNA1-1) has been recently reported to function in autophagy [40].
- **Y RNA** - Y RNAs are ca. 100 nucleotides long Pol III transcripts[41] that associate with the RNA-binding protein Ro60. Ro60 is an autoantigen [42] and functions in RNA quality control by binding misfolded RNA molecules [43].
- **Alu elements** - Alu element are a special class of short interspersed nuclear elements (SINEs) that are Pol III-transcribed retrotransposons with a length of ca. 300 nucleotides [44, 45]. In humans, up to one million copies of Alu elements can exist, which accounts to ca. 10% of the human genome [46]. Alu elements are replicated within the human genome via Pol III-mediated transcription, followed by reverse-transcription and genome integration of the cDNA [47].
- **Brain cytoplasmic 200 long-noncoding RNA (BC200 RNA)** - The BC200 RNA (in primates) or BC1 (in rodents) is ca. 160 nucleotides long and is a brain-specific Pol III transcript [48]. BC200 was suggested to play a role in regulating protein synthesis at the synapse [49] but is also upregulated in cancer cells [50].

1.2.2. Pol III gene structures

Pol III target genes can be classified into type 1, type 2, and type 3 genes (reviewed in [51]). Most of the Pol III genes contain internal control regions (ICRs) that are required for recruitment of Pol III transcription factors. Pol III genes contain a termination sequence that consists of a continuous stretch of thymine (dT) bases on the NT-DNA strand (reviewed in [52]). The length of the dT-stretch, however, can vary between species. The promoter sequences, ICRs, and required transcription factors varies between the three Pol III gene types as described below:

- **Type 1** - The 5S rRNA gene is characterized as a type 1 gene. It contains three characteristic elements: an A-box, an intermediate element (IE) and a C-box [53]. Multiple transcription factors are required to recruit Pol III to the 5S rRNA gene: the 6-subunit TFIIIC complex, the 3-subunit TFIIIB complex (Brf1, Bdp1, TBP), and the single-subunit TFIIIA factor. TFIIIB binds upstream of the ICR elements, TFIIIC binds the A-box and B-box, and TFIIIA binds the IE and the C-box [54]. Whereas TFIIIB and TFIIIC are also required for some of the other Pol III genes, TFIIIA specifically functions in the transcription of the type 1 5S rRNA gene [55].
- **Type 2** - Type 2 gene comprise most of the reported Pol III target genes (tRNAs, VA, Alu, vRNA, EBER, 7SL, BC200). They contain two ICRs, the A-box and B-Box, which are both ca. 10 base pairs (bps) long and separated by a ca. 30-40 bps linker [56]. The 5'-flanking regions of the type 2 genes are diverse and, depending on the gene and organism, may also contain a TATA box (e.g. in tRNA genes of the the silkworm *Bombyx mori* [57]). Type 2 genes require both TFIIIB and TFIIIC, which binds the A-box and the B-box [58].
- **Type 3** - The most prominent type 3 gene is the U6 snRNA gene that contains a strong TATA motif [59]. Depending on the species, U6 snRNA genes may also contain ICRs such as weak A- and B boxes (e.g. in yeast) [60]. Vertebrates contain a characteristic proximal sequence element (PSE) and further upstream the distal sequence element (DSE) [61]. Interestingly, yeast only requires TFIIIB to transcribe the U6 snRNA *in vitro*, whereas animals rely on an additional transcription factor: the snRNA activating protein complex (SNAPc) [62]. SNAPc is made out of 5 subunits in *H. sapiens* (SNAPc1-SNAPc5) and binds the PSE, whereby which it recruits other factors to the target gene such as components of the TFIIIB complex [63]. Notably, the Pol II-transcribed snRNA genes (e.g. U2 snRNA) in *H. sapiens* also contain the PSE and require SNAPc [63, 64]. The specificity between Pol II and Pol III-transcribed snRNAs is given by the presence of the TATA-box in the U6 snRNA gene that is absent in the Pol II-transcribed snRNA genes [63, 65]. The

H. sapiens genome further encodes the protein BRF2, which is also a component of TFIIIB and is homologous to BRF1 (Brf1 in yeast) [66]. Human TFIIIB can either contain BRF1 (used e.g. for type 2 genes) or BRF2 (used e.g. for type 3 genes) [66].

1.2.3. Pol III structure, subunits and regulation

With 17 subunits, Pol III is the largest eukaryotic RNAP (Fig. 1.2 e). The core of yeast Pol III consists of: C160, C128, AC40, AC19, ABC27, ABC23, ABC14.5, ABC10α, ABC10β, and C11 (reviewed in [67]). The five ABC subunits are shared between Pol I, Pol II, and Pol III, and the AC40 and AC19 subunits are shared between Pol I and Pol III. The two largest subunits - C160 and C128 - contain all the structural key elements described in Section 1.1.4. The stalk is formed by C25 and C17 [68]. Pol III also features two Pol III-specific subcomplexes: the C53-C37 heterodimer [69] and the C82-C34-C31 heterotrimer, which was first identified in human cells [70]. While Pol I also contains a heterodimer (A49-A34.5 [71, 72]), the heterotrimer is unique to Pol III. Interestingly, the Pol III-specific subcomplexes share homology to some of the general transcription factors (TFs) in the Pol II system [73].

The Pol III heterodimer: The C53-C37 heterodimer is homologous to TFIIF, which is also heterodimeric and build out of the protein TFIIFα and TFIIFβ. In both the Pol III heterodimer and in TFIIF, the two subunits dimerize via a dimerization module (DM) and bind to the lobe domain of RPC2 (Pol III) and RPB2 (Pol II) [13, 16]. In the Pol II system, TFIIF functions in transcription initiation (reviewed in [74]). TFIIF first binds Pol II, and the Pol II-TFIIF complex gets subsequently recruited to the promoter DNA to form the core PIC. The C53-C37 heterodimer functions both in transcription initiation and termination of Pol III [69, 75–77]. Photo-cross-linking experiments showed that C53-C37 not only interacts with the Pol III core and heterotrimer subcomplex but also with the Pol III transcription initiation factors TFIIIC and TFIIIB [76]. The cryo-EM structure of the yeast Pol III PIC on the U6 snRNA gene further showed that C37 interacts with the TFIIIB subunit Bdp1 [77]. In contrast to its homologs TFIIFα and A49 (Pol I), the yeast heterodimer does not contain a winged-helix (WH) domain that would bind the promoter DNA in context of the Pol III PIC. C53-C37 also interacts with the N-terminal domain of the Pol III subunit C11 as shown by the first cryo-EM structure of yeast Pol III [16]. In *S. cerevisiae*, the C53 subunit has been shown to be hyperphosphorylated by the kinase Mck1, which repressed tRNA and ribosome synthesis under cellular stress conditions [78]. Furthermore, C53 is a hotspot for the

post-translation modification with the small ubiquitin-like modifier (SUMO), shown by genetic screens in *S. cerevisiae*, which represses Pol III activity [79]. The majority of PTM sites in C53 (119-232 for phosphorylation [78] and 51-236 for sumoylation [79]), however, lie in the N-terminal region of C53 (1-247), which is disordered and not visible in the yeast Pol III structure [16].

The Pol III heterotrimer: The C82-C34-C31 heterotrimer also shares homology to general TFs in the Pol II system: TFIIE and TFIIF [73, 77, 80]. The largest subunit, C82, consists of four WH domains and a C-terminal coiled-coil (CC) domain [16, 81]. The first WH domain is homologous to the N-terminal WH domain in TFIIFβ. Together with the other WH domains and the CC domain, C82 binds the Pol III core (primarily the clamp domain of C160) and the other heterotrimer subunits C34 and C31 [16, 81]. The C34 subunit contains three WH domains, of which only the third was resolved in the yeast Pol III apo and EC structures [16]. A comparison between the yeast PIC structure of Pol III [77] and Pol II [13] showed that the structure and position of C82-C34 resembles the dimerization interface and position of TFIIEα-TFIIEβ. Furthermore, the Pol III PIC structures revealed that the first and second WH domains of C34 become ordered and contact the transcription initiation factor TFIIIB. Similarly to the WH-domains of TFIIFα and TFIIEβ, the WH1 and WH2 domains of C34 bind the upstream promoter DNA, albeit at different positions [77, 80]. Hence, the structural data indicate that the Pol III heterotrimer is TFIIE- and TFIIF-like. In line with these observations are biochemical studies, which showed that the Pol III heterotrimer functions in transcription initiation by recruiting Pol III to the promoter DNA and participate in promoter DNA opening [70, 82, 83]. The smallest heterotrimer subunit is subunit C31, of which only a short helical segment could be resolved in the first Pol III cryo-EM structure, where it binds the CC-domain of C82 [16]. Still, additional cryo-EM density between the heterotrimer and Pol III stalk domain could also be assigned to C31, which suggested that C31 links these two modules [16]. More recent work on the Pol III PIC [80], further suggested that the same bridging area of C31 folds into a helix in context of the Pol III PIC, and this region was therefore assigned as the 'stalk bridge'. C31 also participates in transcription initiation as deleting the C-terminus of C31 showed a growth phenotype at restrictive temperatures and failed to initiate transcription *in vitro* despite the presence of Pol III TFs [84].

The RNA-cleaving subunit C11: Another Pol III subunit that shares homology with a TF in Pol II is the subunit C11. C11 is composed of an N-terminal Zinc ribbon domain and a C-terminal Zinc ribbon domain that are connected by a linker [16, 85]. While the

N-terminal domain (NTD) is homologous to the Pol II subunit RPB9, the C-terminal domain (CTD) is homologous to domain III of the Pol II transcription factor TFIIS that also harbours a zinc ribbon domain [85]. Notably, Pol I also contains a homologous subunit, A12.2, which also resembles a hybrid of RPB9 and TFIIS [17]. C11, A12.2, and TFIIS all possess a hydrolytic 3' to 5' RNA cleaving activities, for which they utilize an acidic hairpin (D91, E92 in yeast C11) [85, 86]. Hydrolytic RNA-cleaving activity of Pol III could first be shown by Whitehall et al. [87], which was later assigned to the C11 subunit [85]. The yeast Pol III cryo-EM structure showed that the C11 NTD binds the Pol III core at similar positions as RPB9 in Pol II and the NTD of A12.2 in Pol I [16]. Interestingly, the C11 CTD was only resolved in the apo Pol III structure where the acidic hairpin points towards the Pol III heterotrimer. In sharp contrast, the CTD of A12.2 (Pol I) is inserted into the RNAP funnel in the crystal structures of yeast Pol I in the apo state [17, 88]. Thereby, the position of the A12.2 CTD resembles the one of domain III of TFIIS in the structure of yeast Pol II in its 'reactivation intermediate' conformation [89]. In the EC structures of both Pol I and Pol III from *S. cerevisiae*, the CTDs of A12.2 and C11 got rearranged and were not visible [16, 90]. In the yeast Pol III structure, the C11 NTD also binds the C53-C37 heterodimer. This interaction is critical for the stable association of the C53-C37 heterodimer to the Pol III core because in *S. cerevisiae* - substitution of endogenous C11 with *Schizosaccharomyces pombe* is viable but results in loss of the C53-C37 from the Pol III core during the purification of Pol III [75]. The resulting Pol III Δ variant that lacks C53, C37, and C11 has been used for several subsequent studies to characterize the role of these three proteins biochemically (e.g. see [91–93]). C11 plays an important role in Pol III transcription termination as outlined below in more detail.

Pol III interaction partners and regulatory proteins: Various factors interact with Pol III to regulate its activity. To facilitate transcription initiation, Pol III interacts with its general transcription factors TFIIIA, TFIIIB, TFIIIC, and SNAPc as described in Section 1.2.2. Interestingly, the number of transcription factors that are required for transcription initiation of Pol III is greatly reduced compared to Pol II. For the formation of the human Pol II PIC, up to 64 additional polypeptides are necessary (reviewed in [94]) whereas the number of required proteins for the Pol III PIC ranges between 3- (yeast U6 snRNA gene, 8- (human U6 snRNA gene), 9- (e.g. tRNA genes), and 10 polypeptides (5S rRNA gene) [51]. Together with the stable integration of TF-like subcomplexes (TFIIF-like C53-C37 and TFIIE/F-like C82-C34-C31), the strong reduction of required transcription initiation factors resemble a Pol III-specific specialization to cope with the high demands of its target genes (reviewed in [67]).

Pol III activity is further modulated by regulatory proteins, which is particularly important during cell growth and nutrient starvation as the synthesis of Pol III transcripts consumes a large amount of metabolic energy (reviewed in [95]). The major regulator of Pol III is the protein Maf1, which represses Pol III activity as part of the target of rapamycin (TOR) signaling pathway [96]. Maf1 has a molecular mass of 45 kDa, and its activity to repress Pol III is phosphoregulated. Under ideal growth conditions, yeast Maf1 is is phosphorylated by the protein kinase A (PKA), which renders Maf1 inactive because the protein is located to cytoplasm when phosphorylated [97]. Dephosphorylation of Maf1 (e.g. upon nutrient deprivation) is facilitated by the phosphatases PP2A and PP4, upon which Maf1 can be shuttled to the nucleus to repress Pol III activity [98, 99]. A recent cryo-EM structure of yeast Pol III bound by Maf1 shed light on the molecular basis of Maf1-mediated repression of Pol III [100]. Maf1 binds the Pol III clamp domain of C160 and the WH2 domain of subunit C34. The binding site of Maf1 overlaps with the one of the TFIIIB complex. Thereby, Maf1 blocks the active site, competes with TFIIIB, and interferes with transcription initiation.

Pol III activity is further regulated by additional proteins. The Kns1 and Mck1 kinases phosphorylate the C53-subunit of Pol III and, thereby, represses Pol III activity downstream of the TORC1 pathway [78]. The transcription activity of Pol III is also repressed by the tumour suppressors p53 [101] and the retinoblastoma protein (Rb) [102]. Both p53 and Rb primarily target TFIIIB, upon which the activity of Pol III is repressed [103, 104]. Pol III can also be activated by proto-oncogenes such as c-Myc [105] and ERK [106].

1.2.4. Pol III in health and disease

With the synthesis of tRNAs and other short and essential RNAs, Pol III plays a central role in the eukaryotic cell. Not surprisingly, misregulation of Pol III activity is associated with various diseases (reviewed in [107]). The repression of Pol III by tumour suppressors and its activation by proto-oncogenes (described above) suggest an implication of Pol III in tumorigenesis. Pol III, its transcription factors TFIIIB and TFIIIC, and various Pol III target genes are also upregulated in different forms of cancer [108–111], and an increase of Pol III transcription has been shown to be required for malignant transformation [112]. Furthermore, proteins levels of Pol III subunits or Pol III TFs are elevated during tumour transformation [113] and overexpression of the Pol III subunit RPC7α in IMR90 fibroblasts enhanced tumour transformation [114]. The cumulative findings that link Pol III upregulation to tumour transformation point towards a direct role of Pol III in cancer formation, which renders Pol III a potential drug target to combat tumorigenesis.

Pol III has also been shown to play a role in aging [115]. The reduction of Pol III protein levels or inhibiting Pol III activities in *S. cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* extended the lifespan of these model organisms. The authors concluded that Pol III limits eukaryotic longevity downstream of the TOR complex 1 (TORC1) [115]. Hence, the specific inhibition of human Pol III could, potentially, increase the life expectancy of humans.

Pol III malfunctioning is, however, also associated with several rare diseases (reviewed in [52, 107]). Mutations in the human Pol III codings genes *POLR3A, POLR3B, POLR1C, POLR1D, RPC10* (subunit RPC1, RPC2, RPAC1, RPAC2, RPC10) are linked to genetic diseases such as the neurodegenerative disorder hypomyelinating leukodystrophy (HLD) [116–119], the progeroid syndrome (mimicking physiological aging) Wiedemann–Rautenstrauch syndrome (WDRTS) [120], and the developmental disorder Treacher Collins syndrome (TCS) [121]. The reported mutations can, most likely, be attributed to a reduction of Pol III activity as a consequence of protein misfolding, complex assembly, or direct interference with enzymatic activity. Hence, inhibiting Pol III activity would, presumably, not help to cure these diseases or to improve the living condition of an individual suffering from such a disease. Still a better understanding of the molecular basis of Pol III malfunctioning as a consequence of these genetic mutations could, in the future, aid gene therapy approaches to treat these rare genetic diseases.

Pol III also plays a role in the innate immune system where it senses viral DNA and, thereupon, activates the RIG-I pathway to induce a cellular autoimmune response [122, 123]. Inhibition of Pol III also promoted intracellular growth of bacteria [123]. Mutations in *POLR3A, POLR3C, and POLR3E* (subunit RPC1, RPC3, RPC5) are also associated with an increased vulnerability to infections by the varicella zoster virus (VZV) [124, 125]. These observations can, most likely, also be attributed to a defect of Pol III functioning as part of the innate immune system.

1.2.5. The human Pol III system

The majority of structural studies has been conducted on yeast Pol III [16, 77, 80, 100, 126], which limits our mechanistic understanding of the human Pol III transcription system. The only structures that were available at the beginning of my PhD project were the crystal structures of truncated RPC3-RPC6 (C82-C34 in yeast) at 2.8 Å [81] and a low resolution crystal structure of RPC3 with a short fragment of RPC7 (C31 in yeast) at 7.0 Å resolution [127]. Similarly to yeast Pol III, human Pol III consists of 17 subunits that show a varying degree of sequence conservation between the two species (Table 1 and Fig. 1.5). The highest sequence identities can be observed in the Pol III core with

subunit RPC2 being the most conserved Pol III-specific subunit (ca. 62% identical to its yeast ortholog C128). The peripheral subunits - in particular those of the Pol III-specific subcomplexes - are less conserved (18.18% to 27.12%).

Table 1. Subunits of *S. cerevisiae* (Sc) and *H. sapiens* (Hs) Pol III. Listed are the protein names, molecular weights in kDa, and the respective sequence identities between the two species (according to the percentage identity matrix created by multiple sequence alignment generation in Clustal 2.1 [128]).

Protein (Sc)	**Size (Sc) [kDa]**	**Protein (Hs)**	**Size (Hs) [kDa]**	**Identity (Sc vs Hs) [%]**
C160	162	RPC1	155	50.3
C128	129	RPC2	127	62.3
C82	74	RPC3	61	18.5
C53	47	RPC4	44	20.6
C37	32	RPC5	80	18.2
C34	36	RPC6	36	23.7
C31	28	RPC7	26	27.1
C25	24	RPC8	23	43.7
C17	19	RPC9	17	30.3
C11	13	RPC10	12	51.3
AC40	38	RPAC1	39	45.3
AC19	16	RPAC2	15	44.6
ABC27	25	RPABC1	25	43.5
ABC23	18	RPABC2	14	53.4
ABC14.5	17	RPABC3	17	35.7
ABC10a	8	RPABC4	7	39.0
ABC10b	8	RPABC5	8	73.5

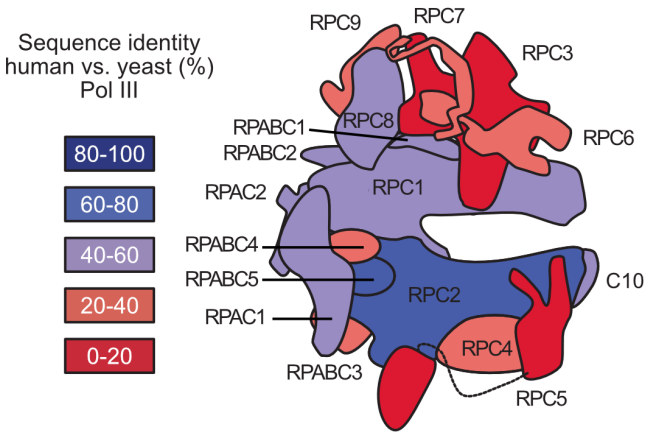

Fig. 1.5. Sequence conservation between human and yeast Pol III. Shown is a schematic of the human Pol III complex and its subunits coloured according to the sequence identities between the two species (see Table 1).

The presence of 17 subunits in both human and yeast Pol III indicate that the overall architecture is conserved but the low sequence identities of the peripheral subunits hint towards sequence-specific variations between yeast and human Pol III. Below, I am listing the Pol III subunits that feature some interesting differences between the well characterized yeast Pol III and less understood human Pol III:

RPC5: With 18.2%, subunit RPC5 is the least conserved Pol III subunit. Interestingly, the size of RPC5 largely varies across species. With a molecular mass of 80 kDa, human RPC5 is more than twice as large as its yeast ortholog C37. Both orthologs feature the N-terminal DM domain, with which RPC5 (C37) dimerizes with RPC4 (C53). The differences in polypeptide length can be attributed to a long C-terminal extension in human RPC5 [129], of which no structural or functional information is available.

RPC6: RPC6 (C34 in yeast) is one of the heterotrimer subunits. Human RPC6 has been shown to contain a cubane iron-sulphur (4Fe-4S or FeS) cluster, which is coordinated by a C-terminal FeS binding domain of RPC6 [130]. Surprisingly, the four cysteins that coordinate the FeS cluster have all been lost in *S. cerevisiae*. Because all available Pol III structures are derived from *S. cerevisiae*, the structure, function, and position of the FeS cluster is not known.

RPC7: RPC7 (C31 in yeast) is the smallest heterotrimer subunit. Human cells can express two variants of RPC7: RPC7α and RPC7β [114]. Whereas the RPC7β variant is ubiquitously expressed, expression of the RPC7α variant has been shown to be stronger in non-differentiated stem cells and tumour cells [114, 131–133]. Overexpression of RPC7α - but not of RPC7β - has further been shown to enhance tumour transformation of IMR90 cells [114]. Furthermore, downregulation of RPC7α reduced proliferation of human prostate cancer PC-3 cells whereas downregulation of RPC7β did not affect proliferation of PC-3 cells. Hence, human cells can express a RPC7 variant that is associated with tumorigenesis but the molecular basis of how RPC7 variant incorporation may affect Pol III activity and tumorigenesis is not known. Because *S. cerevisiae* only encodes one form of the RPC7 ortholog C31, the structures of yeast Pol III also do not contribute to a better understanding of how RPC7 variants selection may modulate Pol III activity.

1.2.6. Pol III transcription termination

Pol III is highly specialized for the production of its target genes, and a key specialization is transcription termination. In yeast, Pol III terminates on sequences that contain a continuous stretch of at least 5 dT bases on the NT-strand of the DNA (Fig. 1.6 a). This process does not rely on external protein factors, which is in sharp contrast to the other eukaryotic RNAPs that rely on additional pausing and terminating factors to trigger termination (reviewed in [134]. Instead, Pol III solely uses its in-built protein factors to terminate, which allows a fast recognition of the termination signal and subsequent release of the DNA and the nascent RNA chain. This process, together with the short distance between the transcription start site and the termination site, enables efficient recycling of Pol III so that it can start a new transcription cycle. Moreover, Pol III can undergo a process called facilitated recycling, in which the transcription re-initiation rate is much higher than in the initial round of transcription [135]. Facilitated recycling also requires termination on the canonical poly-dT termination sequence. A later study showed that the C11 subunit is critical for facilitated recycling of Pol III, which, however, does not rely on its RNA cleaving activity [69].

The key protein elements that participate in Pol III transcription termination are the subunits C160, C128, C53-C37, and C11. Pol III has first been demonstrated to terminate on its own in *in vitro* transcription assays using purified Pol III from *X. laevis* [136, 137], which was shortly after also demonstrated for mammalian [138] and yeast Pol III [139]. The first Pol III subunit that was demonstrated to be essential for Pol III termination was C128 using genetic screens combined with termination reporter assays [140–142]. The next protein that has been shown to participate in Pol III termination was subunit C11 because an *S. cerevisiae* Pol III variant that lacked C11 failed to terminate transcription *in vitro* [85]. Using a similar protein preparation strategy, in which endogenous *S. cerevisiae* C11 was replaced with its ortholog from *S. pombe*, a Pol III variant was purified that lacks C11 and the C53-C37 heterodimer subcomplex [69]. This incomplete Pol III variant (termed Pol III Δ) failed to terminate transcription on a 7-dT termination signal, which could be rescued by the addition of recombinant C53-C37. The essential role of the C53-C37 heterodimer and of C11 in Pol III transcription termination could also be validated in *in vivo* termination reporter assays using *S. pombe* as a model system [143]. Similarly designed experiments also showed that the C160 subunit is critical for correct transcription termination [144]. Several C160 point mutations were identified in genetic screens that caused termination read-through *in vivo* and *in vitro*. The *in vitro* experiments revealed that termination-deficient C160 variants are faster, indicating that Pol III pausing is critical for correct termination [144].

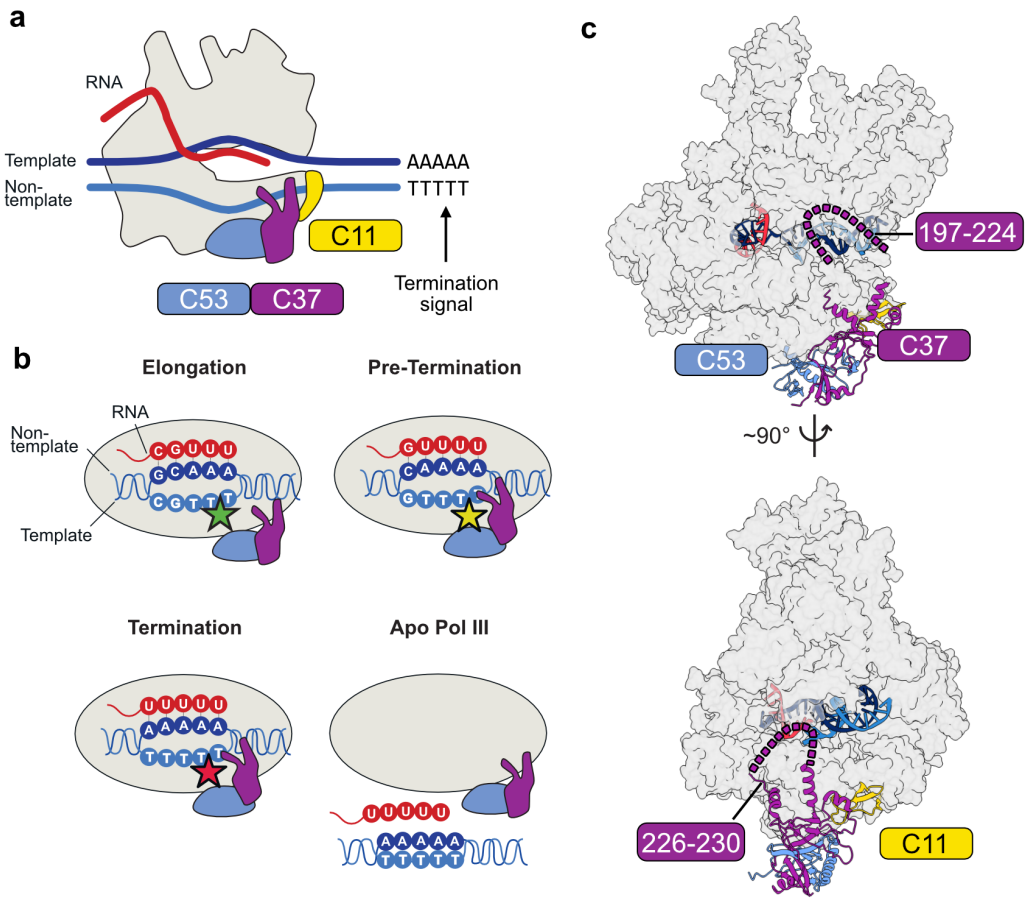

Fig. 1.6. Pol III termination elements. (a) Schematic of elongating Pol III that is about to reach its poly-dT termination signal. **(b)** Schematics illustrating the different expected conformational re-arrangements once Pol III reaches the termination signal. **(c)** Side- and front view of the yeast Pol III elongation complex (PDB: 5FJ8). The Pol III core, stalk, and heterotrimer are shown as a grey surface. C53, C37, and C11 are shown as coloured ribbons. The C37 loop (197-224), which is not resolved in the structure, is illustrated as a dashed line.

A 17 Å 3D reconstruction of yeast Pol III obtained via cryo-EM placed the C53-C37 heterodimer close to the Pol III cleft and in proximity to the downstream DNA where C53-C37 could, in principle, sense the termination signal [145]. A more detailed *in vitro* analysis of the function of C53-C37 and C11 revealed that Pol III is slowed down when approaching its termination site, which is facilitated by C53-C37 and is critical for correct termination on 6-dT stretches [91]. The Pol III Δ variant can also terminate transcription but requires at least 8 dTs, indicating that Pol III termination is facilitated by multiple processes that likely work hand-in-hand. A follow-up study by the same group showed that, when Pol III reaches the 4th dT on the termination signal, it forms a metastable pre-termination complex (PTC) [92] (Fig. 1.6 b). PTC formation requires the C37 subunit and is concomitant to Pol III pausing prior to termination, which hints

towards conformational changes that occur when Pol III adopts the PTC conformation. A direct sensing of the poly-dT termination signal by C37 was suggested to induce these conformational changes that also prime Pol III for subsequent release of the RNA transcript. This hypothesis was further strengthened when the first high-resolution cryo-EM structure of yeast Pol III was reported [16] (Fig. 1.6 c). The C53-C37 heterodimer primarily binds the Pol III core on the lobe region of subunit C128, and C37 places two helical elements in close position to the DNA binding cleft. The two helices are connected by a 27-residue loop (197-224), which was disordered in both apo and EC Pol III (Fig. 1.6 c). Because the deletion of neighbouring residues (226-230) caused termination read-through [92], the authors hypothesized that the C37 loop directly senses the poly-dT signal when the termination signal is reached and might become ordered when the Pol III PTC forms. The cryo-EM structure of yeast Pol III also showed that the active site of Pol III is more open compared to the Pol II active site [16]. Also, the DNA:RNA hybrid is more loosely bound in the Pol III EC, which gives another hint how Pol III can terminate on short dT-stretches. The bound DNA:RNA hybrid is an important stabilizing element of the RNAP EC [146], and the more loosely-bound hybrid in the Pol III EC suggests that Pol III is more sensitive to instable DNA:RNA hybrids. When Pol III approaches the poly-dT termination signal, a poly-dA:rU hybrid has to form, which is less stable than other forms of DNA:RNA hybrids [147]. Thus, it is likely that pausing of Pol III is primarily achieved by sensing the poly-dT termination signal on the NT-strand, whereas the final trigger of termination and release of the nucleic acids could result from a collapse of the unstable poly-dA:rU hybrid in the active site (also discussed in [148]). However, the precise mechanism of how Pol III pausing and termination is achieved on a molecular basis remains speculative due to a lack of structural information.

Interestingly, UV cross-linking experiments on nascent Pol III transcripts showed that a portion of Pol III enzymes reads through primary weak termination sites (5 to 6 dTs) of tRNA genes [149]. This suggested that Pol III termination can fail on shorter dT termination signals and requires secondary termination sites, located downstream of primary termination signals. A more recent study also identified the DNA/RNA helicase Sen1 as a novel Pol III termination factor in *S. pombe* [150]. Sen1 is best known to participate in transcription termination of Pol II (reviewed in [134]) but Rivosecci et al. showed that deleting Sen1 increased accumulation of Pol III downstream of primary termination sites [150]. Given that Pol III is able to terminate solely on its own *in vitro*, it is likely that Sen1 assists termination on weak termination sites where Pol III would fail to terminate by itself.

1.3. Cryo-electron microscopy

Over the last years, cryo-electron microscopy (cryo-EM) emerged as a powerful tool to study the three-dimensional structures of large macromolecular complexes (reviewed in [151, 152]). Compared to X-ray crystallography, cryo-EM offers the advantage that the macromolecule of interest does not have to be crystallized. This is, particularly for large macromolecular proteins, often the main bottleneck for the structure determination via X-ray crystallography. In contrast, cryo-EM benefits from a larger size of biological molecules because the larger mass aids the detection and visualisation of the molecule of interest.

The small wavelength of high-energy electrons (100-300 kEV) is in the range of picometers, which, in principle, allows structure-determination of the specimen of interest at high-resolution. The structures of proteins and nucleic acids can be visualized using a transmission electron microscope (TEM) because biological molecules scatter electrons when exposed to a direct electron beam. This process creates an image of the molecule of interest, which is subsequently magnified with the electron lenses of the imaging system. In biological cryo-EM, the image is typically magnified at a range of 10,000 - 130,000 times and the image is detected with an electron detector.

To enable high-resolution structure determination of biological macromolecules, the specimen needs to be embedded in a thin layer of vitreous ice (reviewed in [153]). First, the TEM has to be operated at high-vacuum since electrons are also scattered by atoms in the air. The vitreous ice layer serves to protect the specimen from the vacuum. Second, the incident electron beam can damage the biological specimen, which is highly susceptible to radiation damage. When operated at cryogenic temperatures, the vitreous ice layer can fix the protein in a near-native state and can partially protect the specimen from the radiation damage. To obtain a high-resolution structure, it is critical that the ice layer is not too thick because the electrons can otherwise not pass the specimen. In addition, it is important that the specimen is being frozen to cryogenic temperature at a fast rate (ca. 10^6 °C/s) so that a vitreous ice layer forms. Otherwise, crystalline ice forms, which cannot be passed by the electrons. Ice contamination also damages the specimen and can result in poor image quality. To achieve the formation of vitreous ice, the biological sample is, typically, applied to a cryo-EM grid (in most cases made out of copper or gold) and then plunge frozen into liquid ethane [154, 155]. The cryo-EM grids contain a layer of holy carbon, which supports embedding of the molecules in the vitreous ice-layer within the holes.

In single particle cryo-EM, the purified macromolecules are applied to the cryo-EM grid and, when plunge-frozen, ideally adopt different orientations, which can be captured

when recording the images. One recorded image typically contains tens to hundreds images of individual biological particles that represent 2D projections of the 3D object. Averaging and classifying these 2D projections computationally then allows to reconstruct the 3D structure of the macromolecule of interest [156]. The computational workflow of single particle analysis has been heavily optimized over the last 30-40 years [157]. One of the most-widely used programs for the single particle cryo-EM analysis is RELION (REgularized LIkelihood OptimizatioN), which implements a Bayesian approach for the 3D reconstruction of the structure of interest [158]. Importantly, RELION can also be used to process structurally heterogeneous cryo-EM data, and, thereby, different conformations or biological states of the molecule of interest can be captured [159]. Notably, the introduction of direct electron detectors was another key development that enabled and greatly accelerated the high-resolution structure determination of biological macromolecules [160]. Not only did the usage of direct electron detectors improve the signal-to-noise ratio and image contrast but it also enabled the collection of the recorded images as multi-frame movies, which allowed the computational correction of beam-induced motion [161].

1.4. Scope of this thesis

Pol III is a fascinating molecular machine that plays an important role in the eukaryotic cell. Despite recent advances in the structural characterization of yeast Pol III, many questions regarding the structure, function, and regulation of Pol III remain to be answered. In this thesis, I aim to address two key questions in the field of Pol III transcription.

The first question asks how human Pol III is structurally organized and how the human Pol III structure differs from the yeast Pol III structure. As outlined in Section 1.2.5, several interesting, but poorly understood, differences between human and yeast Pol III subunits have been reported. Because these features (the C-terminal extension of RPC5, the FeS domain of RPC6, the different variants of RPC7) don't exist in *S. cerevisiae* Pol III, their molecular role in the human Pol III system remains essentially unknown without a high-resolution structure of human Pol III. Since Pol III plays an important role in human health and disease, obtaining structural insights into the human Pol III is also of biomedical interest because it could contribute to a better understanding of how genetic mutations affect the function of Pol III. Furthermore, a high-resolution structure of human Pol III could foster the development of chemical compounds that specifically target Pol III.

The second question focuses on the mechanism of Pol III transcription termination, which is an important specialization of Pol III to cope with the high demands of its transcripts. Although the capability to terminate on poly-dT termination signals in the absence of any co-factors has been described almost 40 years ago, the underlying molecular mechanism is still not known. To better understand Pol III transcription termination, I am interested in how Pol III is able to specifically recognize the poly-dT termination signal on the NT-strand, and how this induces pausing of Pol III. Furthermore, I want to get insights into the role of the C53-C37 heterodimer and the C11 subunit in transcription termination. Lastly, I aim to characterize the functional and mechanistic link between Pol III pausing and subsequent termination. Of particular interest is the structure of the Pol III PTC, which has been suggested to pause Pol III when reaching the termination signal and primes Pol III for subsequent termination. Because Pol I and Pol II, on the one hand, rely on external factors to terminate and, on the other hand, must not terminate on gene-internal poly-dT stretches, I also hope to better understand how the structural elements that drive Pol III termination differ between the three eukaryotic RNAPs. Because transcriptional pausing is a universal requirement for transcription termination, the here-obtained insights could also aid the understanding of transcription termination by other multi-subunit RNAPs.

Part II.
Results and discussion

2. Structure of human Pol III

This section describes the structural analysis of human Pol III via single particle cryo-EM, which presupposes highly pure and homogeneous biological sample. When I initiated this project, I considered two potential options for obtaining human Pol III:

- **Endogenous human Pol III**, purified from a human cell line containing affinity-tagged Pol III. This approach is comparable to the purification of endogenous yeast Pol III, which is well established in the lab and would, therefore, have the advantage that the purified sample is likely intact and active. Another advantage would be that one could, potentially, purify both human Pol I and Pol III at once by adding an affinity-tag to subunit RPAC1 that is shared between Pol I and Pol III. To insert such an affinity-tag, CRISPR-Cas9 would be a suitable tool to genomically engineer the host cell line. A challenge could, however, be that the obtained yields are not sufficient for cryo-EM, which can, potentially, be addressed by using human suspension cells that can grow to high densities such as the Expi293F cell line.

- **Recombinant human Pol III**, purified from an eukaryotic expression cells that overexpress one or more human Pol III coding genes of which one carries an affinity-tag. This approach would have the advantage that, in case the complex assembles correctly, the yields might be higher than of the endogenous protein purification method because the complex is strongly overexpressed. Another benefit would be that the complex could be easily mutated or truncated for structure-function studies to validate drawn conclusions once the structure has been solved. On the contrary, a challenge might be that the complex is not assembled correctly because certain chaperones or post-translation modifications are required, which can, potentially, be addressed by co-expressing the genes in their closest-possible expression host. Again, human Expi293F cells might be a promising cell line.

Given that both approaches have their own advantages and disadvantages, we decided to split distribution of tasks. Whereas, I was focusing on the recombinant expression strategy, Agata D. Misiaszek, who is working on the structure of human Pol I, worked on the endogenous purification strategy. Both approaches yielded pure and intact human Pol III in sufficient amounts for structural and functional characterization.

2.1. Structural analysis of endogenous human Pol III

2.1.1. Endogenous human Pol III is intact and active

For practical reasons (cost- and time efficiency), I decided to use the endogenous human Pol III, purified and kindly provided by Agata D. Misiaszek, for the structural analysis via cryo-EM. As shown by SDS-PAGE and mass-spectrometry (MS), the endogenous human Pol III sample is pure and contains all 17 Pol III subunits (Fig. 2.1 a). An RNA extension assay, performed by Dr. Florence Baudin, further showed that the purified human Pol III was active because it could extend radioactively-labeled RNA, annealed to an artificial DNA:RNA transcription scaffold upon the addition of NTPs (Fig. 2.1 b). To further assess the sample quality, I performed negative-stain EM, revealing that the purified sample is homogeneous and contains well-separated particles of the expected size of approximately 20 nm (Fig. 2.1 c). Furthermore, the shape of the 3D reconstruction of human Pol III was similar to that of yeast Pol III (Fig. 2.1 d). Overall, these results confirmed that the endogenous human Pol III is well behaved and intact and should thus be suitable for high-resolution structure determination by cryo-EM, with which I proceeded.

2.1.2. Structure of human Pol III

To solve the structure of human Pol III, I first prepared a cryo-EM sample of Pol III in its DNA-unbound state (apo). For the cryo-EM sample preparation, I used the same protocol as for yeast Pol III, which I had already optimized by supplementing the sample with the detergent CHAPSO that has been described to aid cryo-EM sample preparation of bacterial RNA polymerase [163]. The prepared sample showed well distributed particle when I screened the cryo-EM grids of human Pol III apo on a Talos Arctica TEM. Collection and processing of a small dataset yielded 2D class averages and a 3D reconstruction that resembled the expected shape of Pol III (Fig. 2.2).

To obtain high-resolution structures of human Pol III, I collected two datasets on a Titan Krios TEM of human Pol III apo and bound to the same transcription scaffold used for RNA extension assays (Fig. 2.1). The transcription scaffold-bound Pol III resembles the elongation complex (EC). Both datasets yielded high-resolution structures of human Pol III apo and EC after global 3D classification and refinement in RELION [166] (Fig. 2.3). More flexible regions could be resolved via masked 3D classification in RELION (Fig. 2.4). Thereby, I obtained eight cryo-EM maps at 2.8 to 3.4 Å resolution (Fig. 2.5 a,b) that resemble different conformations of human Pol III or showed improved signal for regions that are more flexible: **Map A** - Pol III apo; **Map B** - Pol III EC (termed EC-1);

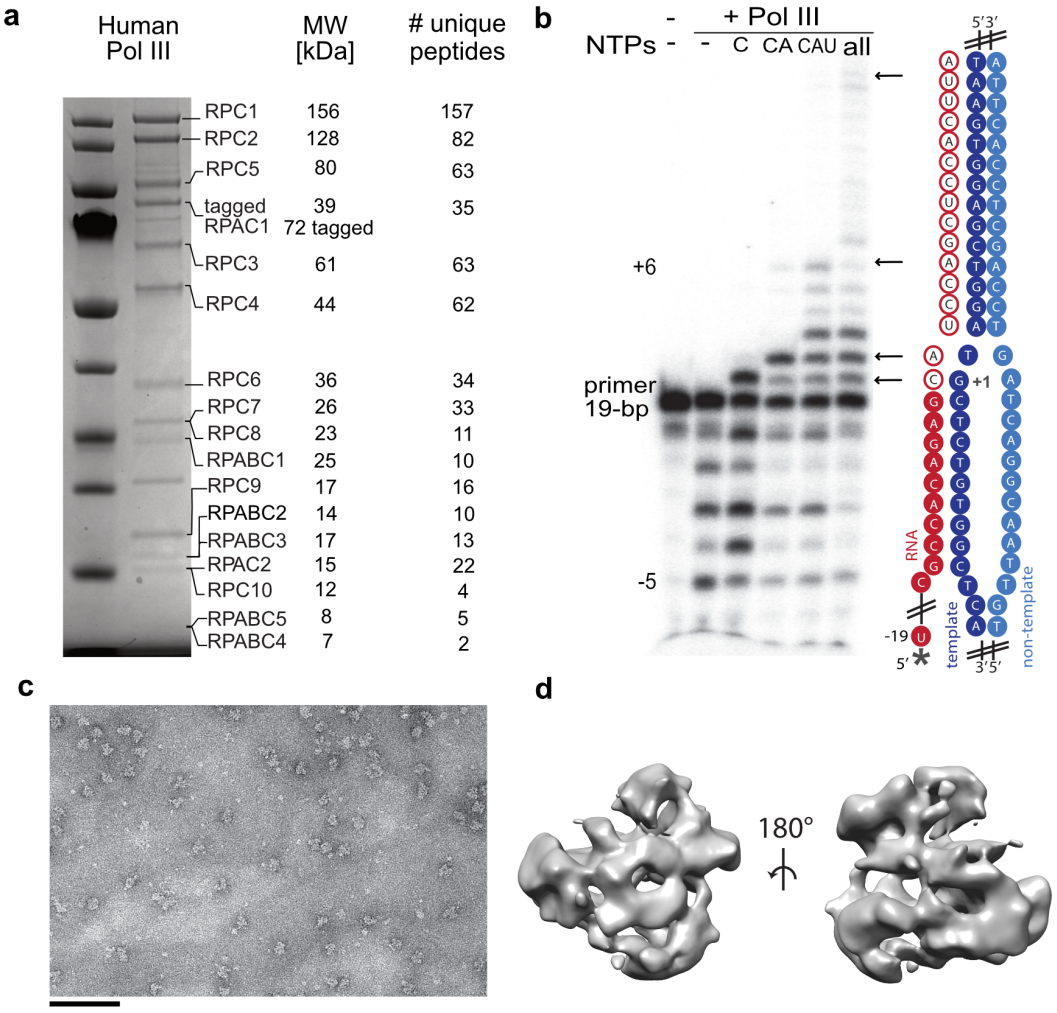

Fig. 2.1. Analysis of endogenous human Pol III by SDS-PAGE, MS and negative-stain EM. **(a)** Coomassie-stained SDS-PAGE (left) and MS analysis (right) of human Pol III. Expected molecular weights (MW) and numbers of uniquely identified peptides are shown. The gel was prepared by Agata D. Misiaszek. MS data were recorded by Per Haberkant and Mandy Rettel (both EMBL Proteomics Core Facility) and analysed by Agata D. Misiaszek. **(b)** RNA extension assay of endogenous human Pol III analysed with a denaturing 17% TBE PAGE. C - CTP; CA - CTP, ATP; CAU - CTP, ATP, UTP; all - CTP, ATP, UTP, GTP. The right side shows the used DNA:RNA transcription scaffold (cartoon representation). The arrows indicate the expected RNA products upon addition of a certain NTP mix. The asterisk marks the radioactively labeled 5'-end of the RNA. The experiment was performed by Dr. Florence Baudin. **(c)** Representative negative stain EM micrograph of human Pol III (scale bar, 100 nm). **(d)** Negative stain EM - 3D reconstruction of human Pol III. Panels **(a)** and **(b)** were prepared by Agata D. Misiaszek. Panels **(a)**, **(b)**, and **(c)** taken from [162].

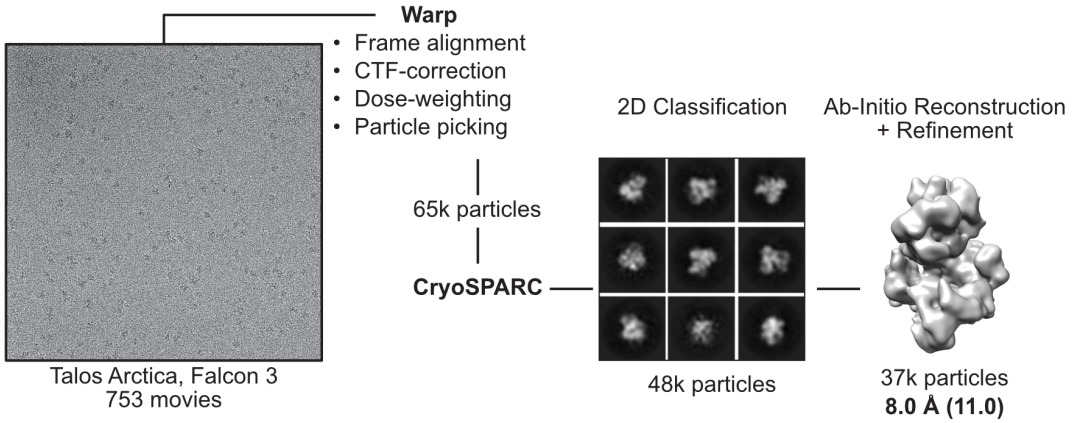

Fig. 2.2. Cryo-EM data collection and processing strategy for a screening dataset of human Pol III apo. (Left): Representative EM micrograph, recorded on Talos Arctica TEM. **(Center)**: selected 2D class averages obtained via 2D classification in cryoSPARC [164] after micrograph pre-processing in Warp [165]. **(Right)**: 3D reconstruction of human Pol III apo reaching a resolution of 8 Å (11 Å unsharpened) obtained via ab-initio reconstruction, followed by homogenous refinement in cryoSPARC. For details of cryo-EM sample preparation and data collection, see Table 4 and Table 5, respectively. Figure taken from [162] in an adapted form.

Map C - Pol III with improved signal for the bound RNA (resembling EC-1); **Map D** - Pol III RPC10 inside funnel domain (termed EC-3); **Map E** - Pol III RPC10 outside funnel domain (resembling EC-1); **Map F** - Pol III with improved signal RPC5 WH 1/2 (resembling EC-1); **Map G** - Pol III with improved signal for RPC4 DM* (resembling EC-1 but was derived from merging the Pol III apo and EC datasets); **Map H** - Pol III with alternative conformation of RPC5 WH 1/2 (termed EC-2).

The Pol III apo map has a nominal resolution of 3.3 Å and, locally, reached 3.0 Å in the core (Fig. 2.5 c). The 3D reconstruction of Pol III EC was resolved at a nominal resolution of 2.8 Å, which extended to 2.5 Å in the core of the polymerase (Fig. 2.5 d). Both maps were of sufficient quality to build the atomic models of human Pol III apo and EC (Fig. 2.6 a,b). Similarly to yeast Pol III [16], the structure of human Pol III can be subdivided into the core (RPC1, RPC2, RPC10, RPAC1, RPAC2, RPABC1-5), stalk (RPC8, RPC9), heterodimer, (RPC4, RPC5) and the heterotrimer (RPC3, RPC6, RPC7). For the Pol III EC, I could also build the downstream DNA that unwinds into non-template (NT) and template strand to which the transcribed RNA hybridises (Fig. 2.6 c). Six bases of the DNA:RNA hybrid in the active center could be build. Akin to the yeast Pol III EC, the NT-strand and the upstream DNA are disordered [16]. Built-in protein elements that are key for eukaryotic RNA polymerase functioning such as the bridge helix, trigger loop and the wall show well-resolved side-chains (Fig. 2.6 d).

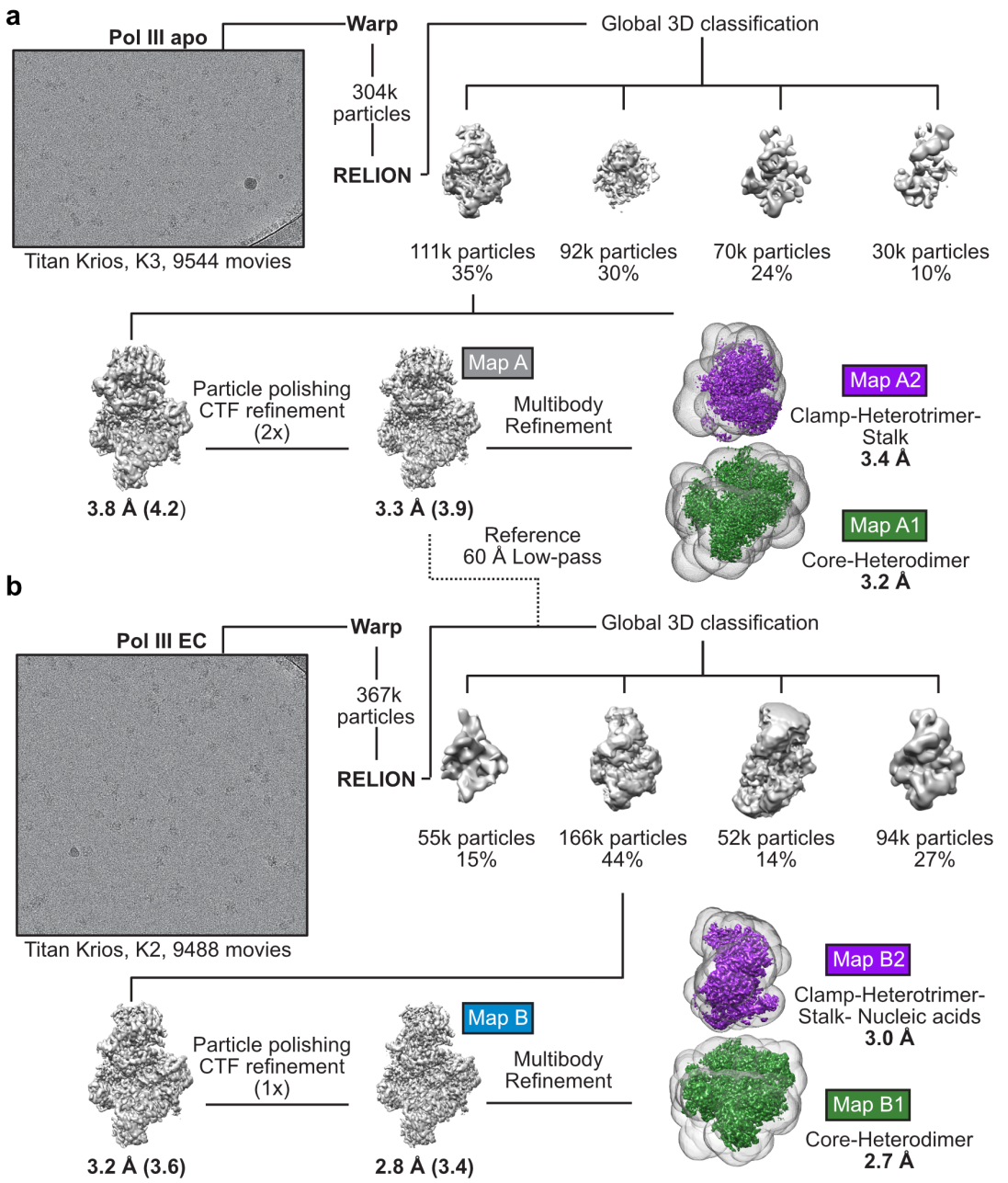

Fig. 2.3. Cryo-EM data collection and processing strategy for high-resolution datasets of human Pol III apo and EC. (a) Processing pipeline of human Pol III apo. For global 3D classification, the Pol III apo map, derived from the Talos Arctica dataset (filtered to 60 Å), was used as a reference model. **(b)** Processing pipeline of human Pol III EC. High-resolution maps were subjected to RELION multi-body refinement [167]. Applied masks are shown as transparent surfaces. Representative micrographs are shown on the left. Reported resolution values are shown below the sharpened and unsharpened maps (in parenthesis). Particle numbers were rounded down. For details of cryo-EM sample preparation and data collection, see Table 4 and Table 5, respectively. Figure taken from [162] in an adapted form.

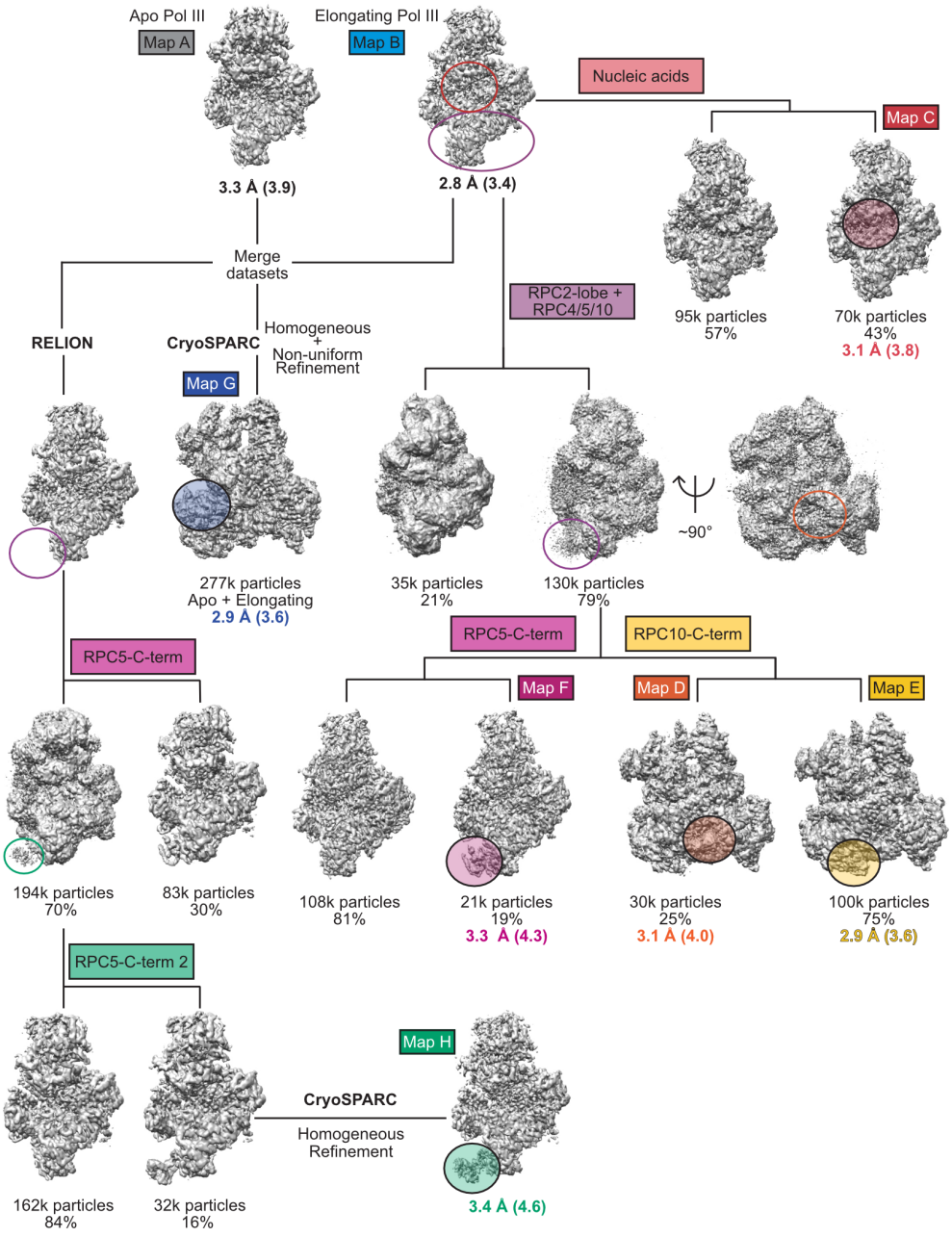

Fig. 2.4. Masked 3D classification strategy to better resolve flexible regions of human Pol III. Coloured unfilled circles mark regions where soft masks were applied and filled, transparent circles show the respective features displaying improved cryo-EM density. Map threshold levels were individually adjusted. The figure is taken from [162].

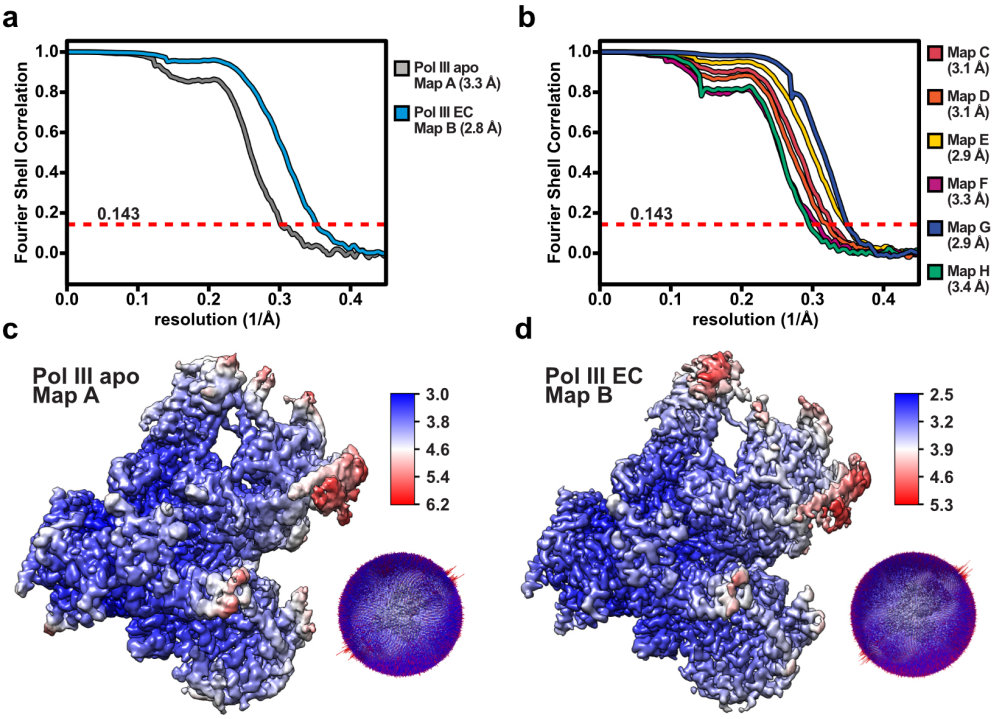

Fig. 2.5. Cryo-EM data resolution assessment (a) FSC curves of map A (Pol III apo) and map B (Pol III EC) **(b)** FSC curves of map C to H obtained via masked 3D classification. Nominal resolution values are given in brackets (FSC = 0.143). **(c)** Local resolution estimation and angular distribution plot (bottom right) of Pol III apo. **(d)** Local resolution estimation and angular distribution plot (bottom right) of Pol III EC. Figure taken from [162] in an adapted form.

Comparing the human and yeast Pol III EC structures reveals that most residues in the human Pol III contacting the nucleic acids are identical or conserved to the contacting amino acids in yeast Pol III (Fig. 2.7 a). Hence, the nucleic acid binding mode is conserved between yeast and human Pol III. In agreement with this, the human and yeast core regions deviate only little from each other (Fig. 2.7 b) as shown by the root-mean-square deviation (RMSD). The deviation of the peripheral regions is, however, more pronounced. I could identify several features in the peripheral Pol III-specific subcomplexes that substantially differ between species, which I describe below.

2.1.3. RPC10 captured inside and outside the Pol III funnel

RPC10 shares homology to the Pol II general transcription factor TFIIS and consists of an N-terminal zinc ribbon (N-ribbon) and a C-terminal zinc ribbon (C-ribbon). A flexible linker connects the two ribbon domains. The C-ribbon features an acidic hairpin (D88, E89) for hydrolytical RNA cleavage on the 3'-end. In both human and yeast Pol III,

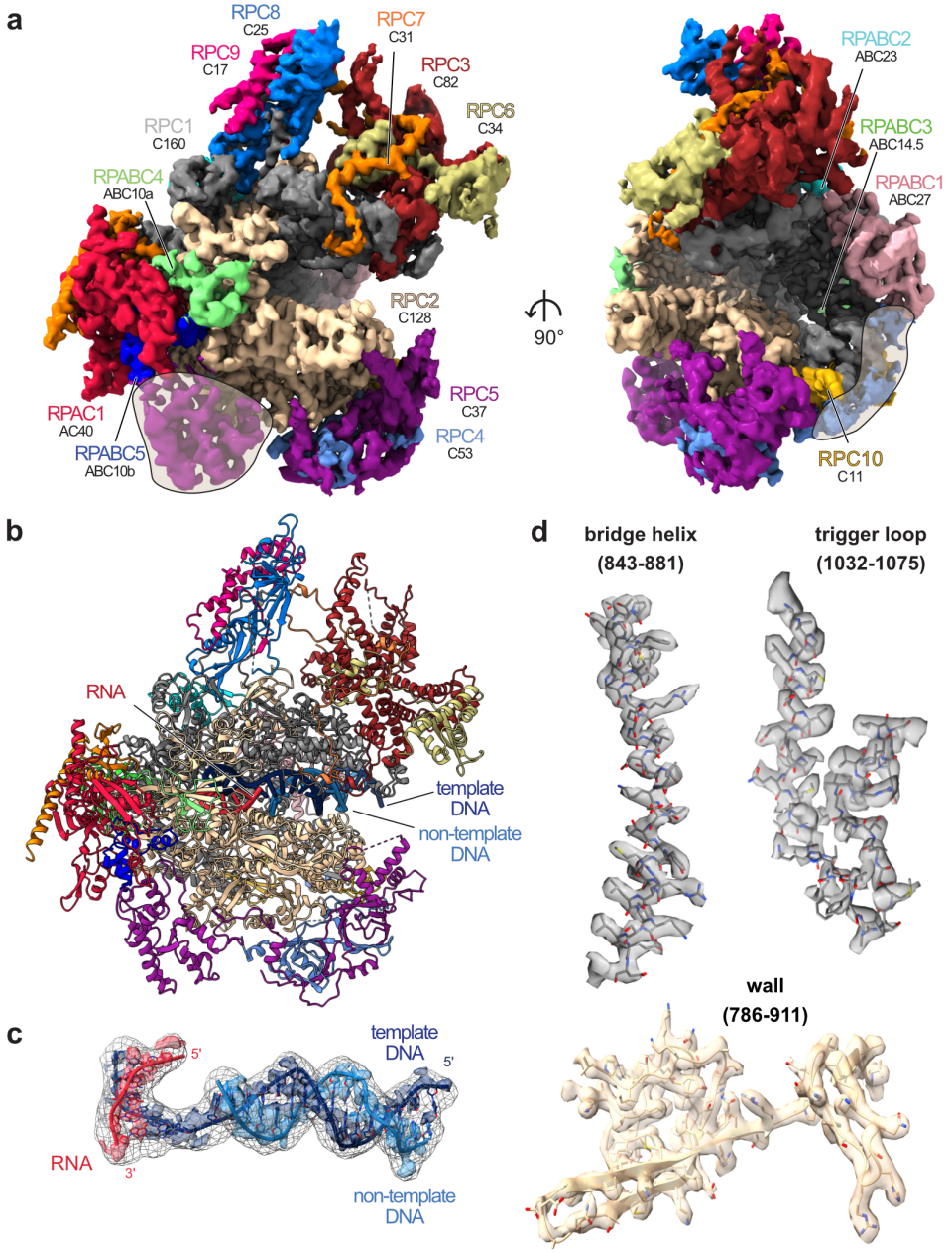

Fig. 2.6. The cryo-EM structures of human Pol III. (a) 3D reconstruction of human Pol III apo (map A) shown from the side and front. The map is coloured according to built-in subunits, which are additionally labeled with their yeast homologs below. The shadings mark superimposed features derived from map G (a pooled dataset of Pol III apo and EC). **(b)** Structural model of human Pol III EC. **(c)** Nucleic acid elements that were built into the cryo-EM density of Pol III EC (coloured surface representation) and overlaid with the same, but low-pass filtered, EM map shown as a mesh. **(d)** Fit of the built-in and labeled active site elements displaying well resolved side-chain densities. Panels **(c)** and **(d)** were prepared by Agata D. Misiaszek. Figure taken from [162] in an adapted form.

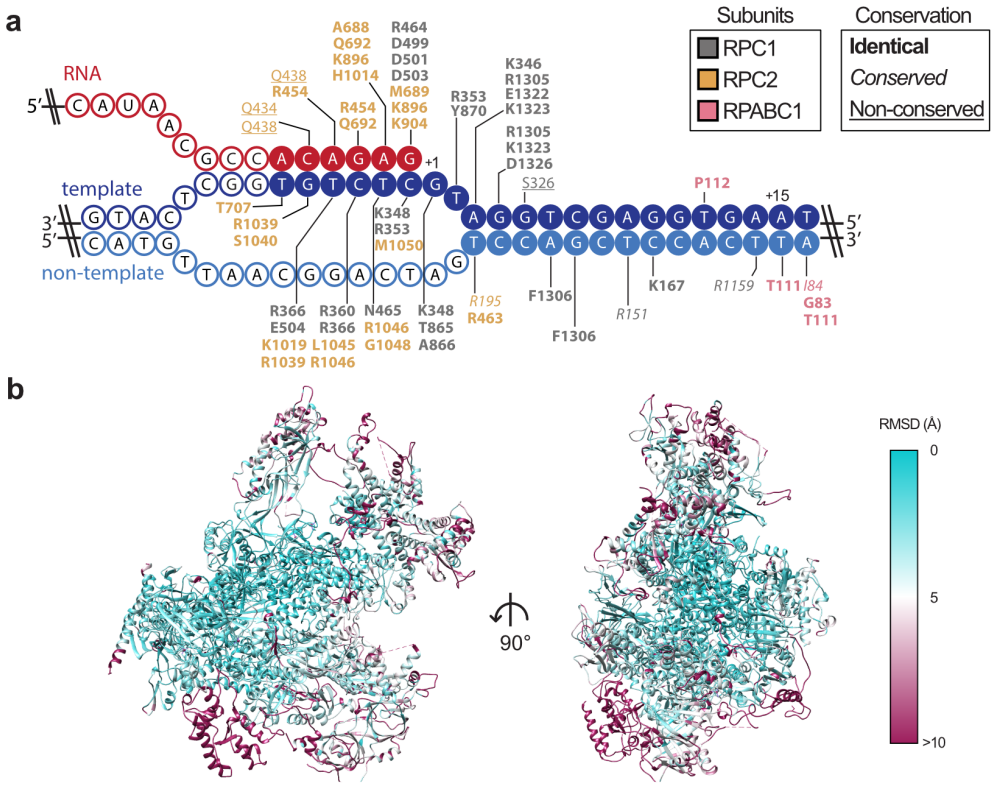

Fig. 2.7. Comparison between human and yeast Pol III. (a) Schematic of the transcription scaffold in the human Pol III EC. Filled and unfilled circles represent modeled and unmodeled nucleotides, respectively. Nucleic-acid contacting residues within a 4 Å distance are indicated. **(b)** Human Pol III EC colored by RMSD to yeast Pol III (PDB: 5FJ8). Nucleic acids are not shown. Newly build regions are colored the same way as RMSD values above 10 Å. Panel **(a)** is taken from [162].

RPC10 (human) and C11 (yeast) are anchored to the Pol III core via their N-ribbons but the C-ribbons adopt different conformations (Fig. 2.8 a,b). In the human Pol III EC-1 structure, the acidic hairpin points towards the Pol III heterodimer and the RPC10 N-ribbon (Fig. 2.8 a). In contrast, the yeast C11 C-ribbon in the Pol III apo complex (the only structure, in which the C11 C-ribbon is visible) points towards the Pol III top where the heterotrimer sits (Fig. 2.8 b). When I built the linker of human RPC10, I noticed continuous density that split-up and either pointed towards the Pol III funnel or the Pol III periphery, which suggested that RPC10 can adopt two conformations. Via masked 3D classification on the Pol III EC map, I could resolve two conformations of the RPC10 C-ribbon, being either outside or inside the Pol III funnel (Fig. 2.4). In the 'outside funnel' conformation (termed EC-1), the linker of RPC10 adopts a kink, and the C-ribbon folds back and binds the linker (Fig. 2.8 a). In the 'inside funnel' conformation (termed EC-3), the linker extends, and the C-ribbon inserts into the Pol III funnel domain (Fig. 2.8 c).

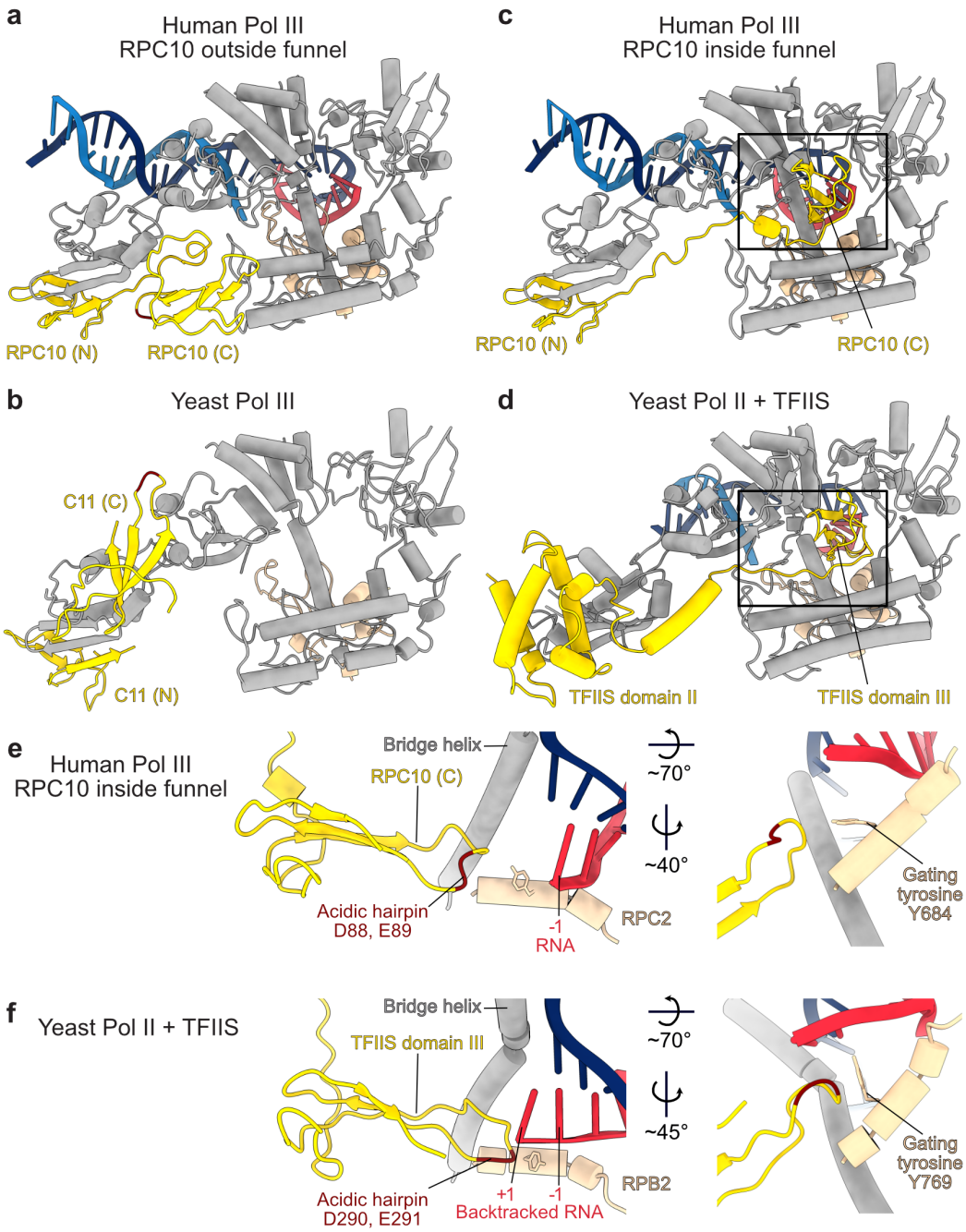

Fig. 2.8. The human RPC10 C-ribbon captured in different conformations. **(a)** RPC10 C-ribbon in the 'outside funnel' conformation. **(b)** The yeast C11 C-ribbon point into a different direction in the yeast Pol III apo structure (PDB: 5FJA). **(c)** RPC10 C-ribbon in the 'inside funnel' conformation in the Pol III EC-3 structure. **(d)** Structure of yeast Pol II in its 'reactivation intermediate' conformation bound by TFIIS (PDB: 3PO3). **(e)** and **(f)** Close-up views of human Pol III RPC10 in the 'inside funnel' conformation and yeast Pol II bound by TFIIS. Gating tyrosines of Pol III RPC2 (Y684) and Pol II RPB2 (Y769) are displayed as sticks. Acidic hairpins of RPC10, C11 and TFIIS are coloured dark red. (N) - N-ribbon; (C) - C-ribbon. Figure taken from [162] in an adapted form.

In the EC-3, the position of the RPC10 C-ribbon closely resembles that of yeast TFIIS bound to Pol II [89] (Fig. 2.8 b). The yeast TFIIS-Pol II structure resembles a 'reactivation intermediate' conformation. In this structure, Pol II exhibits a backtracked RNA molecule, which protrudes out of the funnel and would be cleaved by the acidic hairpin of TFIIS (D290, E291) to 'rescue' backtracked Pol II [89]. The similar positions of the C-ribbons of RPC10 and TFIIS inside the funnels raise the question if the Pol III EC-3 structure also resembles a backtracked Pol III structure. In the human Pol III EC-3, the 3'-RNA end is positioned at the '-1' position, which is a hallmark for an elongating polymerase (Fig. 2.8 e). In the Pol II-TFIIS structure, the RNA is backtracked and the 3'-RNA end is at position '+1' (Fig. 2.8 f). In the study that reported the Pol II reactivation intermediate, a key role was assigned to the gating tyrosine (Y769) in the Pol II subunit RPB2, which needs to flip so that the TFIIS acidic hairpin can reach the RNA for hydrolytic cleavage [89]. In the Pol III EC-2, the gating tyrosine of RPC2 (Y684) is not flipped and shields the RNA. Consequently, the RPC10 C-ribbon cannot reach the 3'-RNA end. Thus, the Pol III EC-3 structure does not reflect a backtracked conformation suggesting that RPC10 can insert into the Pol III funnel during transcription elongation to intrinsically monitor the active site. To get further insights into the conformational dynamics of RPC10, I subjected the 3D reconstruction of the Pol III EC (map B) to the 3D variability analysis (3DVA) tool [168]. The 3DVA revealed that the funnel insertion of the RPC10 C-ribbon is accompanied by a partial opening of the Pol III clamp domain (Fig. 2.9). Both the active-site monitoring activity of RPC10 and the partial opening of the Pol III clamp may explain the role of RPC10 in transcription termination, which will be discussed later in more depth.

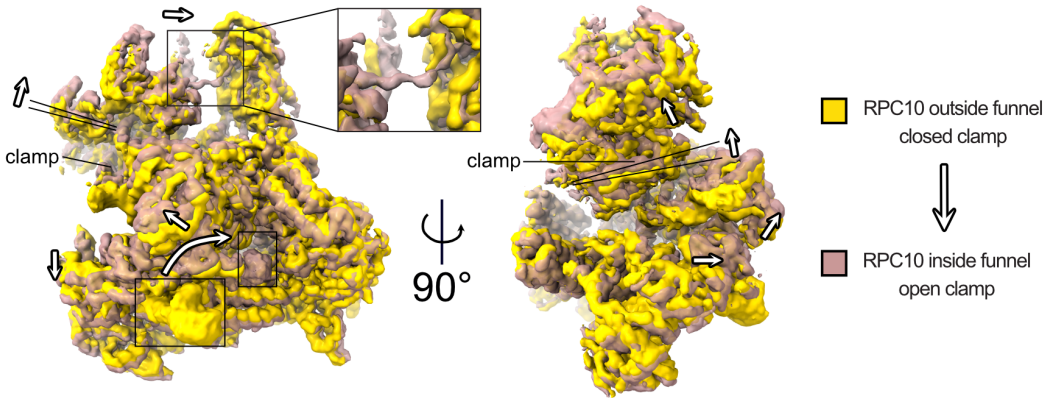

Fig. 2.9. Conformational dynamics of human Pol III. Side and front view of human Pol III EC in two conformations obtained via 3D variability analysis. Yellow map - RPC10 outside funnel; Light red map - RPC10 inside funnel. Boxed regions mark positions of the RPC10 C-ribbon and a close-up view of the connection between the heterotrimer and stalk domain. White arrows mark conformational rearrangements between the two conformations. Figure taken from [162] in an adapted form.

2.1.4. Structure of the human RPC4-RPC5 heterodimer

The human Pol III heterodimer consists of subunit RPC4 and RPC5 that bind each other via their dimerization modules (DM) (Fig. 2.10 a). Human RPC5 (80 kDa) is more than twice as large as yeast C37 (32 kDa). The C-terminal half of RPC5 is predicted to encode four winged-helix (WH) domain. I could place and build the first two WH domains into the Pol III map F, which I retrieved via masked 3D classification (Fig. 2.4, Fig. 2.10 b). The WH1 and WH2 domain are closely intertwined and thus can be assumed to function as a rigid module (WH1-2) (Fig. 2.10 c,d). WH1-2 docks onto the Pol III core close to subunit RPC2, RPAC1 and RPABC5 and is connected to the RPC5 DM domain with a flexible linker (Fig. 2.10 c,d). Using a second masked 3D classification step, I could also obtain an alternative conformation of Pol III (termed Pol III EC-2), in which the RPC5 linker bends and the RPC5 WH1-2 module tilts away from the Pol III core (Fig. 2.10 e).

To get insights into the potential function of the RPC5 C-terminal WH extensions, I initiated a collaboration with Dr. Aleix Lafita and Dr. Alex Bateman (EMBL-EBI, Hinxton, UK) who performed a phylogenomic analysis of the RPC5 WH domains and of the Pol III transcription factors BRF1, BRF2 and the 5-subunit SNAP complex (SNAPc). The phylogenomic analysis revealed that the number of RPC5 WH domains not only differ in their numbers between *H. sapiens* and *S. cerevisiae* but is also diverged across other species (Fig. 2.11 a). When comparing Fungi, Protostomia (e.g. *C. elegans* and *D. melanogaster*) and Deuterostomia (e.g. *G. gallus* and *H. sapiens*), it becomes apparent that Fungi not only lack all RPC5 WH domains but also the two transcription factors BRF2 and SNAPc, which are both required for U6 spliceosomal RNA (U6 snRNA) transcription in *H. sapiens.* [51]. Protostomia appear to contain the WH1-2 module and SNAPc but lack WH3-4 as well as BRF2 whereas Deuterostomia and most plants contain all four WH domains, BRF2 and SNAPc. The comparison led me to hypothesise that the WH1-2 module potentially interacts with SNAPc and that the WH3-4 module specifically binds BRF2 to fine-tune U6 snRNA transcription. In agreement with this hypothesis, modeling of human U6 snRNA PIC indicates that the WH1-2 module in the conformation 2 likely points towards the downstream DNA, to which SNAPc would bind (Fig. 2.11 b). Still, some distant species (e.g. *L. major*) contain all four WH domains but neither BRF2, nor SNAPc, suggesting that the proposed role of the WH domains holds not true for all species. Hence, biochemical and structural analysis is required to better understand the role of the RPC5 WH domains.

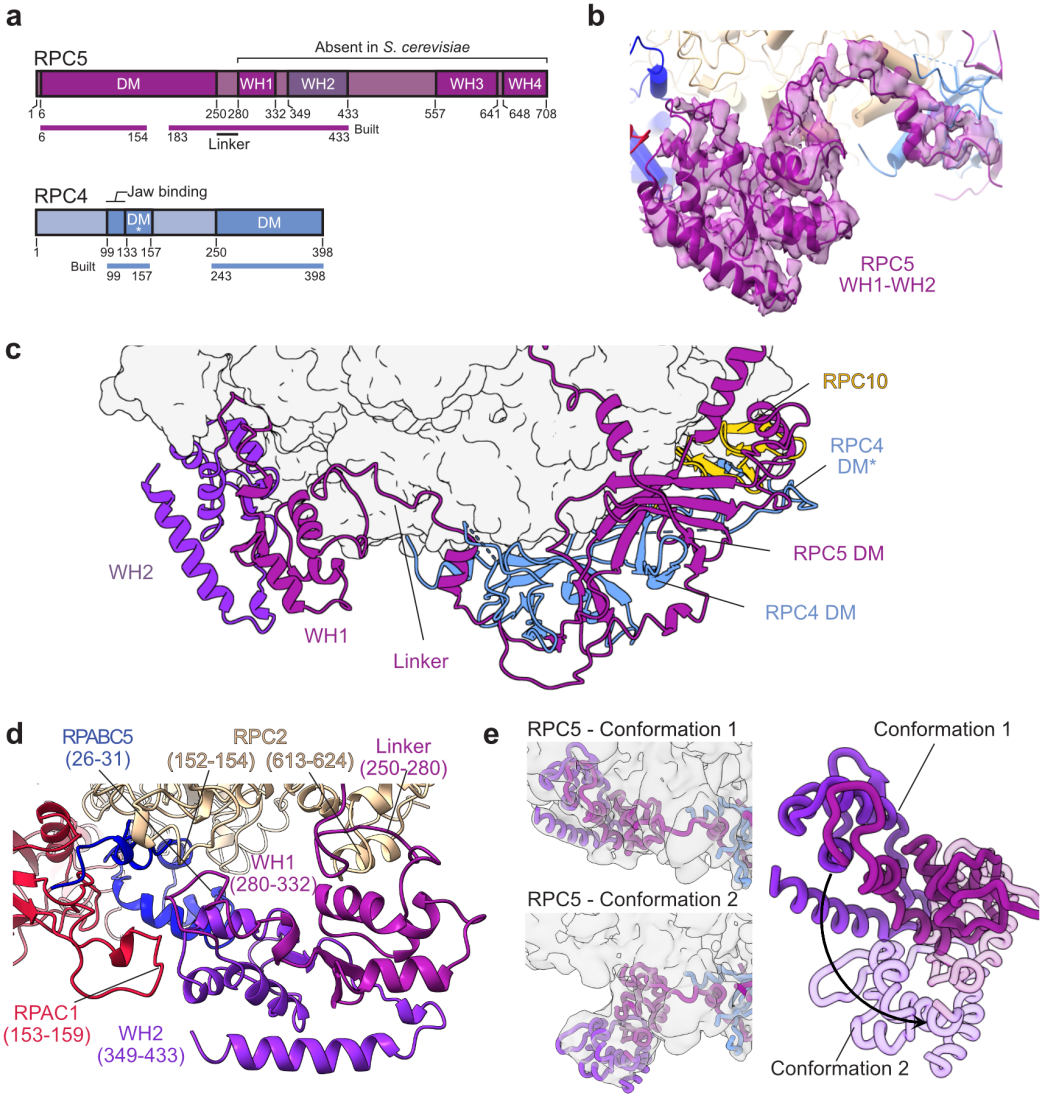

Fig. 2.10. Structure of the human RPC4-RPC5 heterodimer. **(a)** Domain organization of human RPC4 and RPC5. Built regions are labeled with colored bars. The C-terminal half of RPC5 is predicted to encode four WH domains that are absent in *S. cerevisiae* Pol III as indicated. DM - dimerization module; DM* - extended dimerization module; WH - winged helix domain. **(b)** Cryo-EM density fit of the RPC5 WH1 and WH2 domains built into the Pol III map F. **(c)** Structure of human Pol III showing subunit RPC4-RPC5 and RPC10. The built RPC5 model features two WH domains that dock onto the Pol III core and are connected to the RPC5 DM via a flexible linker. **(d)** Close-up view of the newly built RPC5 WH1-2 module binding subunits RPC2, RPAC1, and RPABC5 in Pol III with contacting regions being labelled. **(e)** Left: Different conformations of RPC5 WH1-2 (drawn as thick ribbons): conformation 1 (map F, top) and conformation 2 (map H, bottom). Cryo-EM maps are shown as transparent surfaces and were filtered to 6 Å. Right: Superimposition of the WH1-2 domain in the two different conformations. Panel **(b)** was prepared by Agata D. Misiaszek. Figure taken from [162] in an adapted form.

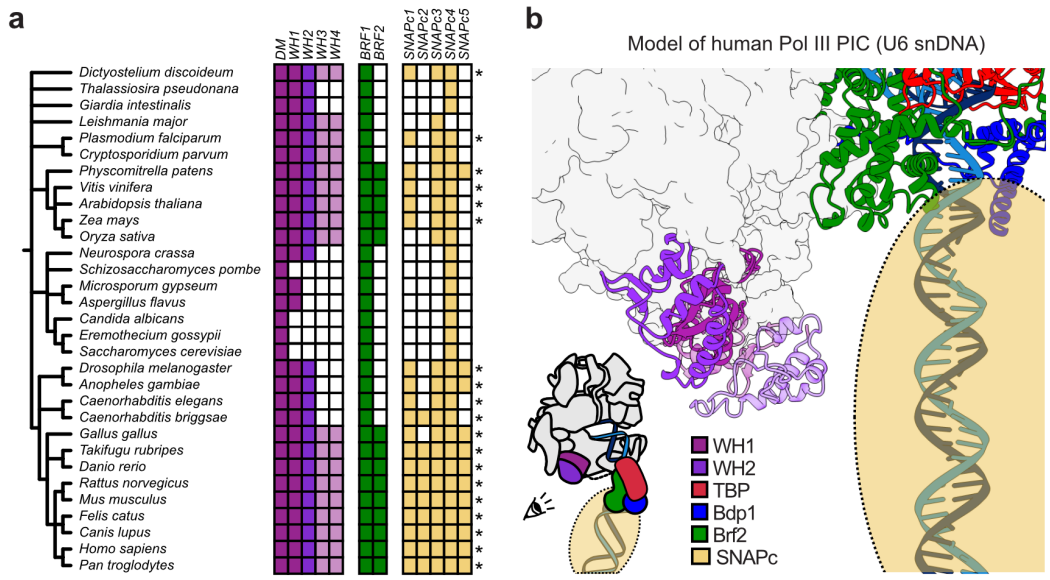

Fig. 2.11. Phylogenomic analysis of RPC5 and Pol III transcription factors and potential role of the RPC5 WH domains in U6 snRNA transcription. (a) Phylogenomic analysis of RPC5, BRF1, BRF2, and SNAPc. Coloured and white squares indicate identified and non-identified homologs, respectively. The asterisks mark species, in which SNAPc1, SNAPc3, SNAPc4 are found, which are required to build a functional complex in *D. melanogaster*. Homologous proteins were retrieved using HMMER [169] and Pfam [170] domain definitions whenever applicable. The analysis was performed by Dr. Aleix Lafita. **(b)** Model of the human U6 snRNA Pol III pre-initiation complex (PIC). The model was generated by: i) superimposition of human TBP (PDB: 5N9G) to yeast TBP (PDB: 6F42); ii) extending the DNA in 5N9G with ideal B-DNA in Coot, iii) superimposing the human Pol III EC model on the yeast Pol III PIC structure (PDB: 6F42). Since no structural model of SNAPc is available, it is drawn as a transparent sphere and placed onto the modelled DNA according to [171]. The two conformations of the WH1-2 are shown with WH1-2 in conformation 1 and 2 depicted as solid and transparent ribbon, respectively. Figure taken from [162] in an adapted form.

2.1.5. Extension and modification of human RPC4

I could assign some extra cryo-EM density to residues 99 to 157 of RPC4 (C53 in yeast). This region, termed DM*, runs along the RPC1 jaw domain and ends in the DNA binding cleft (Fig. 2.12 a). The remaining part of RPC4 (1–99) is not visible in the 3D reconstruction but is predicted to bind DNA (Fig. 2.12 b), raising the possibility that it may bind the downstream DNA in the human Pol III PIC. RPC4 can also be sumoylated (Fig. 2.12 b), and sumoylation of C53 (the yeast homolog of RPC4) represses Pol III activity [79]. Although most of the reported SUMO-modification target sites lie in a disordered region (158–249), one target residues (K141) could be mapped to the DM* domain. K141 is buried in the interface between RPC4, RPC5 and RPC10 (Fig. 2.12 c). Hence, this residue can only be targeted for sumoylation if the RPC4–RPC5 dimer unfolds, which may serve as a signal for proteosomal degradation of defective Pol III.

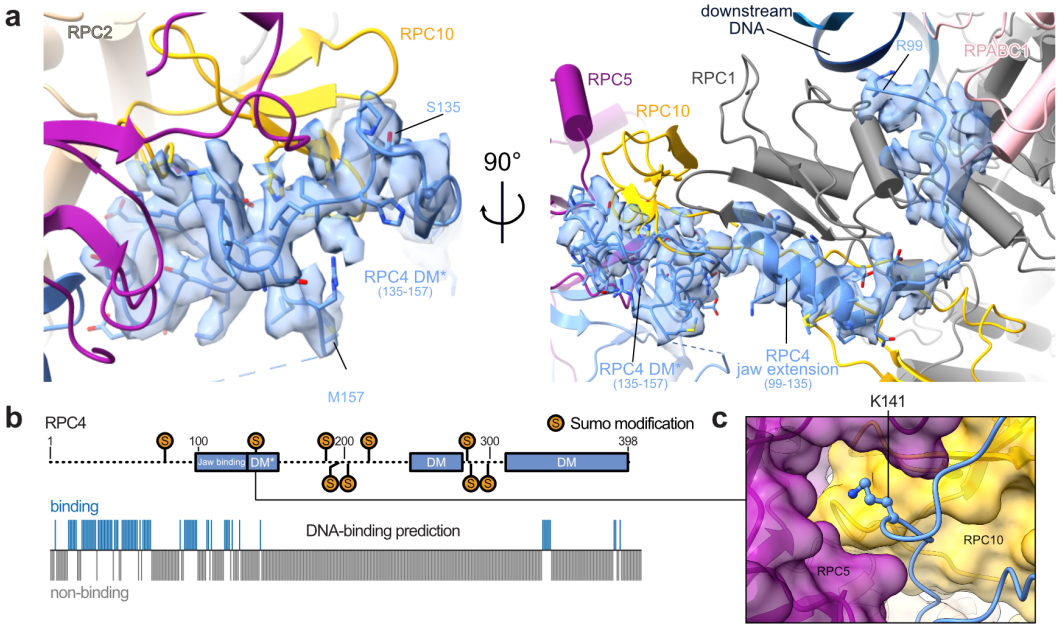

Fig. 2.12. Structure of human RPC4 extending the heterodimer. (a) Cryo-EM density fit of the RPC4 DM* and jaw binding regions built into the Pol III map G. **(b)** Schematic domain organization RPC4. Known Sumo-modifications (K78, K141, K190, K199, K206, K220, K285, K288, K302) are marked and derived from the PhosphoSitePlus database [172]. Per-residue DNA binding prediction of RPC4 is shown below and was obtained from DP-Bind [173]. Dashed lines mark disordered regions. **(c)** Close-up view onto the RPC4 sumoylation target site K141, which is buried within interfaces of RPC4, RPC5 and RPC10 DM. Panel **(a)** was prepared by Agata D. Misiaszek. Figure taken from [162] in an adapted form.

2.1.6. Structure of the human heterotrimer

The human heteotrimer is formed by subunits RPC3, RPC6 and RPC7 (Fig. 2.13 a). The overall shape of the human heteotrimer resembles the one from yeast (made out of C82, C34 and C31) but I could identify additional features that are absent in yeast Pol III (Fig. 2.13). Subunit RPC6 contains a C-terminal iron-sulphur (4Fe–4S, FeS)-binding domain [130]. Because the FeS binding domain is absent in budding yeasts, as indicated by the lack of the four cysteines that coordinate the cubane FeS cluster, (Fig. 2.13 b) no information about its structure, position and function was available based on the yeast Pol III structures [16]. In the human Pol III structure, the cubane FeS cluster is well resolved, and the C-terminal FeS domain adopts a globular fold that is sandwiched between other elements of the heterotrimer (RPC7 and RPC3 WH1 and CC domains) and the Pol III RPC1 core (Fig. 2.13 c,d). The FeS cluster, thereby, serves as an interaction hub that stabilises the heterotrimer and ties it to the Pol III core. The FeS domain overlaps with the visible N-terminal part (31-46) of *S. cerevisiae* C82 (RPC3) in the *S.*

cerevisiae structure where it also binds C160 (RPC1), C34 (RPC6) and C31 (RPC7) (Fig. 2.13 e). This N-terminal extension appears to be budding-yeast specific (Fig. 2.13 f) and likely fulfils the same role as the human RPC6 FeS by serving as an interaction hub. Thus, budding yeast and other eukaryotes have evolved different mechanisms to stabilize the Pol III heterotrimer and to tether it to the Pol III core.

I also observed additional cryo-EM densities in the human heterotrimer that I could assign to the stalk bridge and clamp-binding (CB) elements of RPC7 (Fig. 2.14). The RPC7 stalk bridge is formed by residues Y73 to W99 (Fig. 2.14 a). It protrudes out of the heterotrimer, binds the stalk with its tip and folds back to enter the heterotrimer again. The tip of the stalk bridge is anchored to the stalk with a conserved tyrosine (Y87) forming a cation-π interaction with residue R107 of the stalk subunit RPC8 (Fig. 2.14 a). The stalk bridge is visible in both human Pol III apo and EC, indicating that it moves together with the stalk during the conformational rearrangements upon DNA binding (Fig. 2.14 b). During this transition, stalk and heterotrimer move into opposite directions. Hence, RPC7 ensures that stalk and heterotrimer stay connected with each other, which may aid opening of the promoter DNA during transcription initiation.

In humans, RPC7 occurs in two isoforms: RPC7α (enriched in cancer cells) and RPC7β (ubiquitously expressed in normal cells) [114]. An N-terminal portion (residue 10-33) docks onto the RPC1 clamp coiled-coils, which may provide clues as to how RPC7α incorporation is linked to a cancer-associated Pol III variant. When superimposing the human Pol III structure with the structure of yeast Pol III bound by the Pol III repressor Maf1 [100], the clamp binding (CB) element of RPC7 would clash with Maf1 (Fig. 2.14 c). It, thus, seems likely that the RPC7 CB element must first be replaced before Pol III could be inhibited by human Maf1, which functions as a tumour suppressor [174]. In yeast, binding of Maf1 to Pol III is stabilized via an aromatic stacking interaction between W294 (Pol III - RPC1/C160) and W319 (Maf1) [100]. In the human Pol III structure, RPC7 also encloses RPC1 F288 (the human counterpart to yeast W294) with a hydrophobic surface (Fig. 2.14 d). Similarly to yeast Maf1 W319, a juxtaposed residue in human RPC7α, Y12, potentially forms aromatic stacking with F288. In addition, a putative hydrogen bond between RPC7α Y12 and RPC1 K39 could be formed. The corresponding residue in RPC7β, L12, would neither be able to form the aromatic stacking, nor the hydrogen bond. Based on this observation, I hypothesize that RPC7α might associate more strongly with the Pol III clamp than RPC7β. Hence, Pol III that incorporates RPC7α might be more resistant to Maf1-mediated repression in cancer cells.

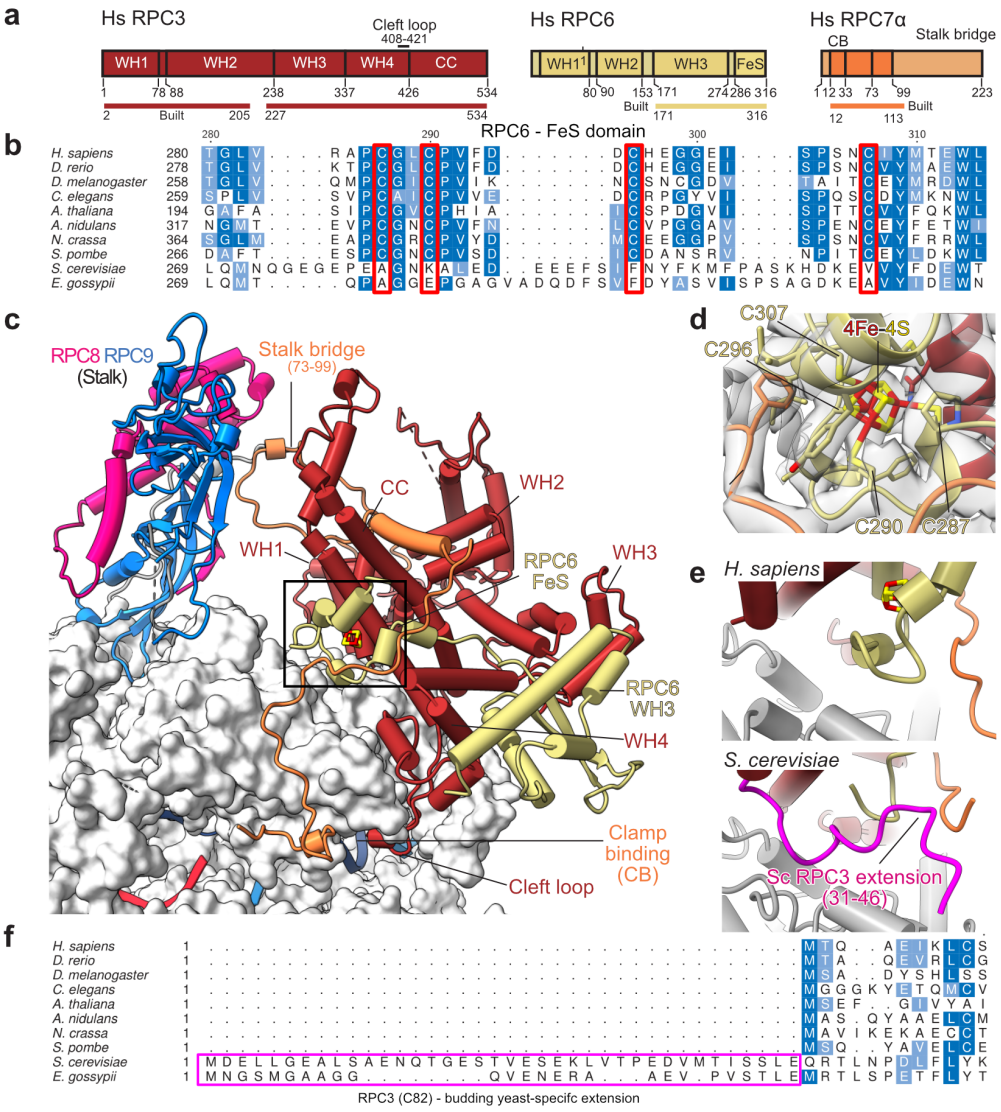

Fig. 2.13. Structure of the human RPC3-RPC6-RPC7 heterotrimer and the RPC6 iron–sulfur binding domain. **(a)** Domain organization of RPC3, RPC6 and RPC7. Built regions are labeled with colored bars. CC - coiled-coil; FeS - iron–sulfur (4Fe–4S) binding; CB - clamp binding. **(b)** Multiple sequence alignment of the RPC6 FeS domain from selected model organisms. The four FeS-coordinating cysteines are highlighted with red boxes. **(c)** Heterotrimer and stalk are shown as cartoons. The Pol III is shown as a white surface. The boxed regions highlights the FeS cluster. **(d)** Close-up view on the RPC6 FeS domain and the cubane FeS cluster. Amino acids within 6 Å distance of the FeS cluster are shown. The four cysteines coordinating the FeS cluster are labeled. The cryo-EM density of map B is shown as a transparent surface. **(e)** Close-up comparison between *H. sapiens* and *S. cerevisiae* Pol III (PDB: 6EU3) heterotrimer-core region. An N-terminal extension of *S. cerevisiae* RPC3 (highlighted in pink) is located at a similar position as the FeS cluster in *H. sapiens* Pol III. **(f)** Multiple sequence alignment of the RPC3 N-termini from selected model organisms. The budding yeast-specific N-terminal extension is highlighted in pink. Figure taken from [162] in an adapted form.

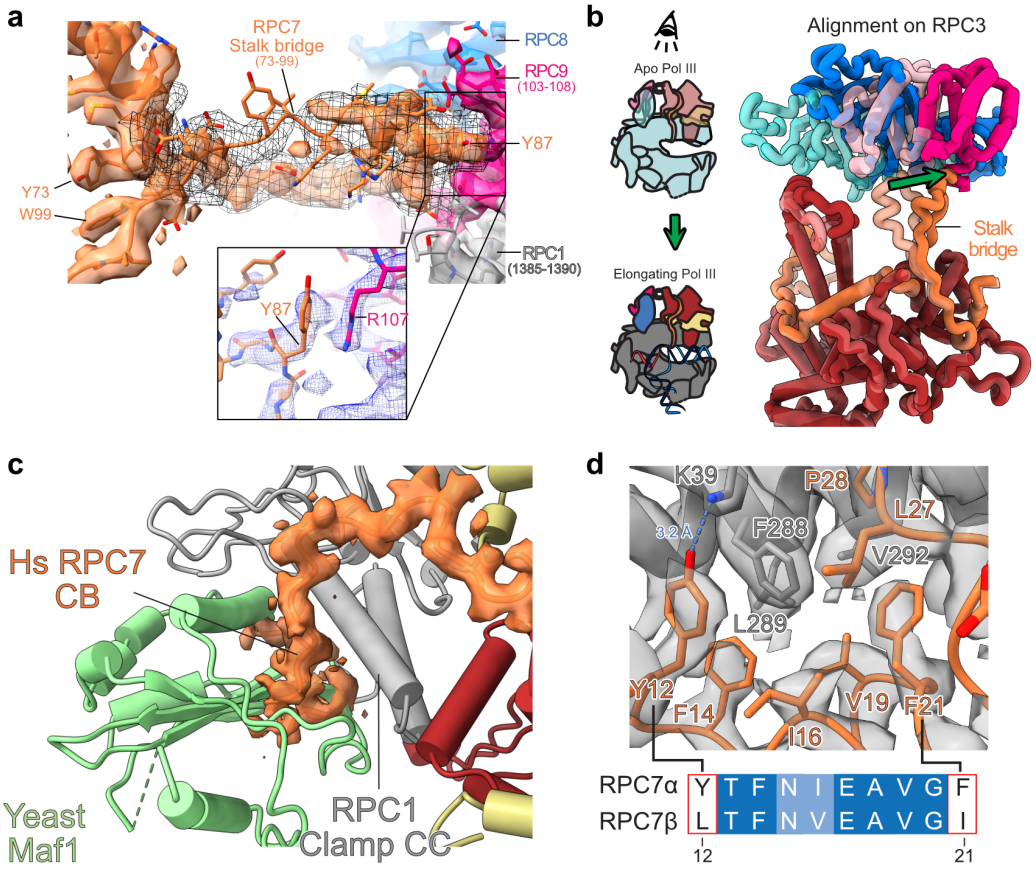

Fig. 2.14. Structural insights into RPC7 as part of the human heterotrimer. (a) RPC7 stalk bridge connects heterotrimer and stalk. Cryo-EM density (map B) corresponding to RPC7 is colored in orange. The overlaid black mesh corresponds to the same cryo-EM map, low-pass filtered to 5 Å. Residues Y73 and W99 that define beginning and the stalk bridge are labeled. The tip of the stalk bridge is shown as a close-up (framed). The blue-mesh corresponds to the cryo-EM map B2, which was obtained via RELION multi-body refinement and sharpened by density modification [175]. **(b)** Top view onto the heterotrimer and stalk domains. Human Pol III apo and EC are shown as transparent and solid ribbons, respectively and were superimposed on RPC3. The RPC7 stalk bridge stays connected to the stalk during the transition from Pol III apo to EC (illustrated on the left and highlighted with a green arrow). **(c)** Superimposition of human Pol III apo and yeast Pol III bound by Maf1 (PDB: 6TUT). Of the yeast Pol III structure, only Maf1 is shown. The cryo-EM density corresponding to RPC7 is shown as a transparent orange surface. **(d)** Top: Close-up view on the interface between RPC7 (orange) and the RPC1 clamp coiled-coils (gray) The cryo-EM density is shown as a gray transparent surface. Bottom: sequence alignment of RPC7α (UniProt: O15318) and RPC7RPC7β (UniProt: Q9BT43). Figure taken from [162] in an adapted form.

2.2. Recombinant human Pol III

2.2.1. Cloning of human Pol III expression constructs

To produce recombinant human Pol III, I first cloned expression constructs for co-expression of all 17 Pol III subunits in Expi293 suspension cells. I used the biGBac cloning system, which combines restriction-free Gibson cloning [176] with restriction-enzyme based release of intermediate constructs for subsequent assembly of large multi-gene expression constructs [177]. The workflow for cloning of the human Pol III expression constructs is summarized in Fig. 2.15.

First, I cloned the coding genes for the 17 Pol III subunits, which were ordered as cDNA clones, into separate pcDNA 3 starting vectors (Fig. 2.15 a). For affinity-purification of Pol III, I added an N-terminal 3x-FLAG-mCherry tag onto subunit RPC2. Second, I assembled the genes into pBig1 intermediate constructs via Gibson cloning (Fig. 2.15 b). The pBig1 constructs contained up to five Pol III coding genes and were designed according to [178] to resemble stable Pol III subassembly modules expected to be generated during Pol III biogenesis inside the cell (Fig. 2.15 c). Third, I joined the pBig1 constructs into a larger pBig2 construct by mixing the pBig1 constructs, digesting them with PmeI and connecting them via Gibson assembly. Initially, I attempted to assemble all 17 Pol III genes into a single pBig2 construct, which was not successful. Hence, I adapted the approach and successfully cloned a 14-gene Pol III pBig2 construct, which contained the subunits resembling the Pol III core (RPC1, RPC2, RPC10, RPAC1, RPAC2, RPABC1-5), stalk (RPC8, RPC9), and heterodimer (RPC4, RPC5) (Fig. 2.15 d, e). The Pol III heterotrimer genes (RPC3, RPC6, RPC7) were kept on a separate pBig1 plasmid. The biGBac system has been designed in such way that all coding genes are separated by a SwaI-restriction enzyme digestion site. The cloned constructs showed the expected digestion patterns (Fig. 2.15 f, g) indicating that the assemblies were successful, which I also confirmed by Sanger sequencing and next-generation full plasmid sequencing.

2.2.2. Purification of recombinant human Pol III

Dr. Helga Grötsch co-transfected Expi293 cells with the P4 and the P6 constructs (Fig. 2.15 e) to express recombinant human Pol III, which I continued to purify (Fig. 2.16). In an initial attempt, I purified recombinant human Pol III from a 60 mL culture using a one-step 3xFlag-affinity purification protocol. The purified protein sample contained SDS-PAGE protein bands at the expected sizes and, overall, showed a band pattern that was similar to the one from yeast Pol III (Fig. 2.16 a). Affinity-purified human Pol III from 330 mL Expi293 cells could be further purified via size-exclusion chromatography

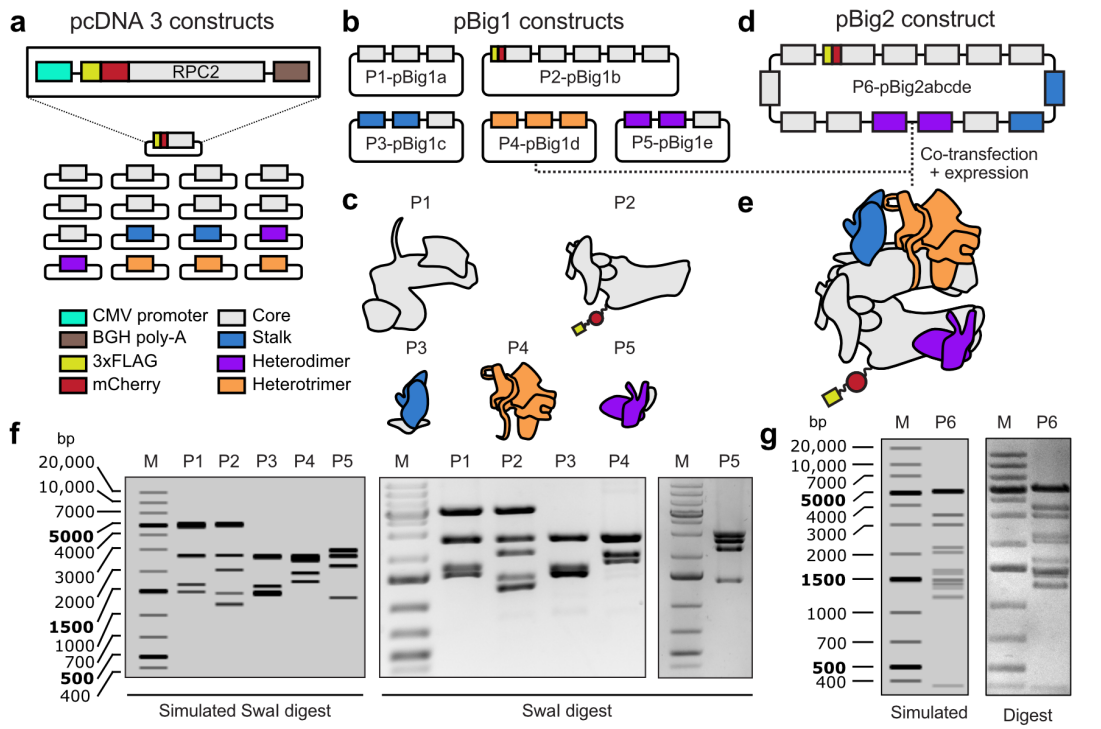

Fig. 2.15. Cloning of human Pol III expression constructs via biGBac. (a) The 17 pcDNA 3 starting constructs. Each gene is under control of a CMV promoter and a BGH-polyA signal as indicated on the top scheme showing affinity-tagged RPC2 as an example. **(b)** Cloned pBig1 constructs resembling modules of the Pol III core, stalk, heterodimer and heterotrimer subcomplexes as illustrated in **(c)**. **(d)** Cloned pBig2 construct generated by assembling the Pol III pBig1a, pBig1b, pBig1c, and pBig1e constructs into pBig2abcde. **(e)** Schematic of the full-length 3x-FLAG-mCherry-labeled human Pol III complex produced by co-transfection of Expi293 cells with the pBig2-core-stalk-heterodimer and the pBig1d-heterotrimer constructs. **(f), (g)** Simulated (left) and real (right) agarose gels (1%) showing SwaI digestion of human Pol III pBig1 constructs P1-P5 **(f)** and of the pBig2abcde construct P6 **(g)**. M - Marker. Gels were cropped for clarity.

(SEC) as shown in Fig. 2.16. However, the purified complex contained a large proportion of co-purified DNA and the SDS-PAGE band pattern indicated that the Pol III complex was non-stoichiometric because the heterotrimer subcomplex was under-represented. After several attempts of optimization, I could overcome both issues by applying the FLAG-purified complex to an anion-exchange chromatography (AEX) column instead of a SEC column (Fig. 2.16 c). AEX separated non-stoichiometric Pol III lacking the heterotrimer (fraction P1) from stoichiometric Pol III (fraction P2). The AEX-purified Pol III complex showed a stoichiometric band pattern (except for a co-purified Hsp70 fraction), which was similar to that of yeast Pol III (Fig. 2.16 d). All 17 subunits were identified by mass spectrometry, and the recombinant human Pol III looked similar to endogenous human Pol III (Fig. 2.16 e). Protein size differences between recombinant

and endogenous Pol III (RPC2, RPAC1) could be explained via different affinity-tagging strategies. In the recombinant Pol III, subunit RPC2 contains an N-terminal 3xFLAG-mCherry tag, whereas in the endogenous Pol III, subunit RPAC1 is C-terminally tagged with an mCherry-StrepII tag.

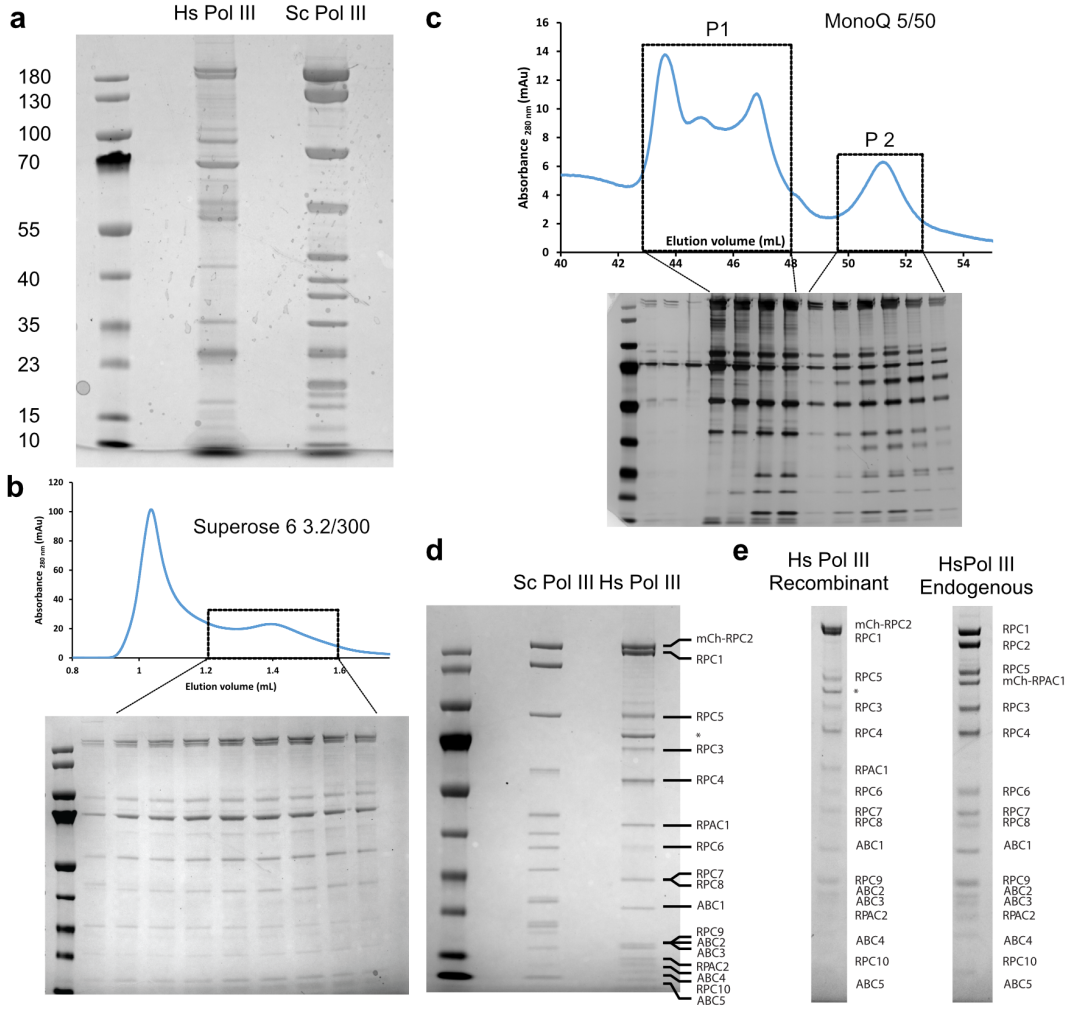

Fig. 2.16. Purification of recombinant human Pol III. (a) Comparison between yeast and human Pol III. Expi293 culture volume: 60 mL; purification: 3xFlag-affinity purification. **(b)** Top: Superose 6 3.2/300 profile of human Pol III. Bottom: SDS-PAGE of indicated SEC fractions. Expi293 culture volume: 330 mL. **(c)** Top: MonoQ 5/50 profile of human Pol III. Bottom: SDS-PAGE of indicated fractions. Expi293 culture volume: 1 L. **(d)** Comparison between yeast and human Pol III (MonoQ-purified). **(e)** Comparison between recombinant and endogenous human Pol III (both MonoQ-purified). Labeled subunits were identified via mass-spectrometry. The asterisk indicates contamination with co-purified Hsp70. Hs - *H. sapiens* (human); Sc - *S. cerevisiae* (yeast). Endogenous Pol III was purified by Agata D. Misiaszek. The gel in panel **(e)** was prepared by Agata D. Misiaszek

2.2.3. Recombinant human Pol III is intact and active

The recombinant human Pol III complex is intact, which I could show via negative stain-EM. Both the 2D classes and the obtained 3D reconstruction resembled the expected shape of Pol III (Fig. 2.17 a,b). Furthermore, *in vitro* transcription assays, performed by Dr. Florence Baudin, showed that recombinant human Pol III synthesises RNA on single-stranded DNA and that its activity is comparable to that of yeast Pol III (Fig. 2.17 c). A direct comparison between recombinant and endogenous human Pol III, also performed by Dr. Florence Baudin, on a tailed-template construct (dsDNA containing a single-stranded template DNA overhang) indicated that the activity of recombinant Pol III is similar to that of endogenous human Pol III (Fig. 2.17 d). Recombinant Pol III was even slightly more active than endogenous human Pol III, which might, however, be due to an underestimation of the used endogenous Pol III protein concentration. Thus, recombinant human Pol III is intact and active. Combined with the modularity of the biGBac cloning system and the fact that Pol III subunits of interest can be easily mutated or truncated, the recombinant human Pol III system can serve as a powerful tool to study the function and regulation of Pol III.

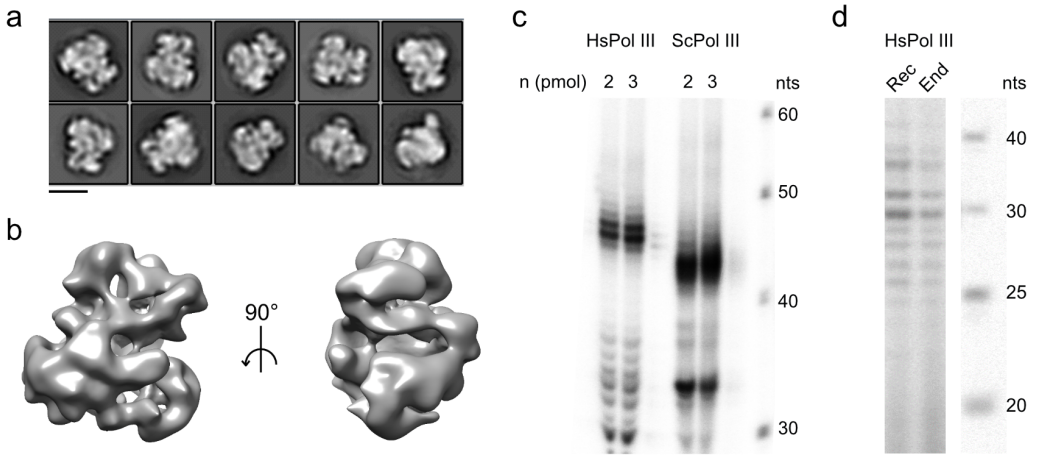

Fig. 2.17. Quality control of recombinant human Pol III via negative-stain EM and *in vitro* transcription assays. (a) Selected 2D class averages of recombinant human Pol III obtained via processing of negative-stain EM data (scale bar 10 nm). **(b)** Negative-stain EM 3D reconstruction of recombinant human Pol III. **(c)** *In vitro* transcription assay on single-stranded DNA (1.3 pmol) comparing recombinant human and endogenous yeast Pol III (2-3 pmol, each). Right lane: DNA-ladder, nts = nucleotides. Hs - *H. sapiens* (human); Sc - *S. cerevisiae* (yeast). **(d)** *In vitro* transcription assay on dsDNA containing a single-stranded template DNA overhang (2 pmol) comparing recombinant (rec) and endogenous (end) human Pol III (2 pmol, each). Right lane: DNA-ladder. Experiments shown in panels **(c)** and **(d)** were performed by Dr. Florence Baudin.

2.2.4. Mutagenesis of recombinant human Pol III for functional studies

The cryo-EM structures of human Pol III gave insights into the commonalities and differences between yeast and human Pol III and also shed light on some general features of Pol III transcription. Based on the structures, I proposed several hypotheses that need to be validated further. The recombinant Pol III is an ideal tool to test these hypotheses as it allows mutating or truncating the areas of interest for the subsequent functional characterization. This subsection revisits the structure-based hypotheses, describes how these ideas can be approached experimentally, and describes the molecular cloning of the constructs to tackle these questions.

Human Pol III heterotrimer variants The cryo-EM structure of the human heterotrimer revealed that human RPC6 features a C-terminal FeS binding domain, which coordinates an FeS cluster and forms a globular fold. The FeS is tightly packed between the other heterotrimer subunits and the Pol III clamp domain, which suggests that the FeS domain plays a critical role in holding the heterotrimer together and tethering it to the Pol III core. To test if the FeS domain indeed has a stabilizing role that is required for Pol III functioning, I designed a truncated version of RPC6 that lacks the FeS, which one could compare with the wild-type (WT) Pol III (Fig. 2.18 a). To rule out any potential contamination of untruncated RPC6 from the endogenous human expression host, I designed two constructs of RPC6, WT (P7) and ΔFeS (P8), which both contain an N-terminal Strep II tag that should allow separating recombinant and endogenous RPC6 from the purified Pol III complex. To test the effect of the FeS deletion on complex stability, one could perform nano differential scanning fluorimetry (nanoDSF). The effect on Pol III transcription could be tested in *in vitro* transcription assays.

Furthermore, the heterotrimer structure gave more insights into structure and function of subunit RPC7. The N-terminal region (2-28) binds the Pol III clamp and would clash with the general Pol III repressor Maf1 when superimposing the human Pol III structure with the one from yeast bound to Maf1 [100]. This suggests that the N-terminal part of RPC7 interferes with Maf1-mediated repression of Pol III. A close-up inspection of the binding interface between RPC7 and Pol III clamp also showed that the cancer-associated Pol III RPC7α isoform could, potentially, interact more strongly with the Pol III core than the ubiquitously RPC7β variant. RPC7α features a tyrosine residue (Y12) that can form a putative H-bond with K39 of RPC1 and an aromatic stacking interaction with F288 of RPC1. RPC7β contains a leucine instead, which would not be able to form these interactions. I hypothesize that a stronger binding of RPC7α to the Pol III core could render such Pol III variant more resistant to repression by Maf1, which could explain

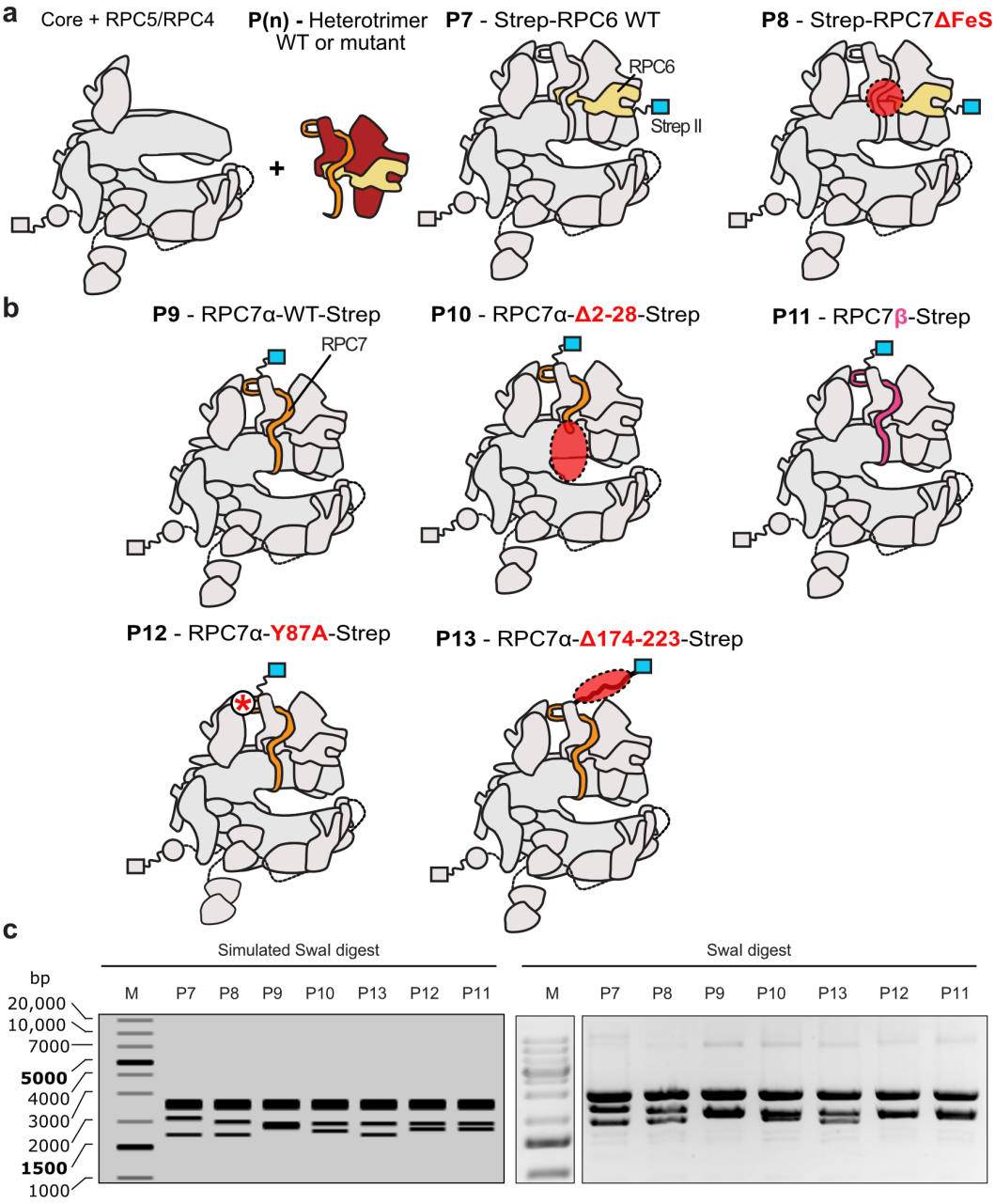

Fig. 2.18. Human Pol III heterotrimer truncation constructs. Schematics of human Pol III variants that were designed to functionally test the hypotheses proposed in this work based on the cryo-EM structures of human Pol III. **(a)** RPC6 variants (WT and ΔFeS) contain an N-terminal Strep II tag. **(b)** RPC7 variants carry a C-terminal Strep II tag. Blue squares: Strep II tag; red ellipses: truncations; red star: mutation. **(c)** Simulated (left) and real (right) agarose gels (1%) showing SwaI digestion of human Pol III heterotrimer truncation constructs P7-P13. M - Marker. Gels were cropped for clarity.

why RPC7α is associated with cancerous cells since Maf1 also functions as a tumour suppressor. This model, however, remains speculative until it can be tested biochemically. I, therefore, designed three RPC7 variants that contain a C-terminal Strep II tag and could be used to test this hypothesis: WT RPC7α (P9), RPC7α Δ2-28 (P10), and WT RPC7β (P11) (Fig. 2.18 b). The effect of these variants on Pol III transcription and Maf1-mediated repression could be tested in *in vitro* transcription assays in the presence or absence of recombinant human Maf1.

Besides the potential role in Maf1-mediated repression of Pol III, I also got insights into how RPC7 can engage with the Pol III stalk subunit and, thereby, functions as a molecular bridge that holds together the heterotrimer and stalk domains. RPC7 interacts with the stalk via a tyrosine (Y87) that forms a cation-π interaction with an arginine (R107) of the stalk subunit RPC9. I propose that this interaction is crucial for connecting the Pol III heterotrimer and the stalk. Since both domains participate in transcription initiation, one could imagine that it is important for Pol III that both domains are connected and that the conformational changes are coupled to each other. Interestingly, the stalk and heterotrimer move into opposite direction upon the "apo" to "elongation" conformation. Since RPC7 remains connected to the stalk in both conformations, the concomitant molecular tension could play a role in promoter DNA opening, for which Pol III requires less external general transcription factors than Pol II. Thus, disrupting the interaction between RPC7 and the stalk domain could give insights into the mechanism of promoter DNA opening by Pol III. I, therefore, designed the RPC7-Y87A (P12) mutation, which should abolish the cation-π interaction between RPC7 and RPC9 (Fig. 2.18 b).

Finally, I designed the truncation RPC7 Δ174-223 (P13), which lacks the C-terminal portion of this subunit (Fig. 2.18 b). I primarily designed this truncation because deleting the C-terminus of RPC7 (C31 in yeast) has been shown to be lethal in *S. cerevisiae* and reduce transcription initiation [84]. In this study, the authors speculated that the C-terminus could interact with the Pol III transcription factors TFIIIB or TFIIIC. Interestingly, it has recently been shown by cryo-EM that the C-terminus of human RPC7 (185–223) also engages with the active site of apo Pol III [179]. Based on this observation, it has been suggested by the authors that human Pol III is autoregulated by the C-terminal tail of RPC7, which first needs to be replaced from the active site and potentially might also interact with TFIIIB. Hence, the designed truncation RPC7 Δ174-223 could be used to test this model in *in vitro* transcription assays.

Human Pol III heterodimer variants The cryo-EM structures of human Pol III gave new insights into the structure and function of the human Pol III heterodimer. Based on

structural and phylogenomic analysis, I propose that the four WH domains of RPC5, which are absent in the yeast ortholog C37, may function in the transcription of the U6 snDNA (Fig. 2.19 a). I hypothesize that the WH domains 1 and 2 interact with the SNAP complex and that the WH domains 3 and 4 interact with the transcription factor BRF2. Both factors are only present in some eukaryotes where they are specifically required for the the production of the U6 snRNA. Furthermore, an extended structure of the RPC4 subunit suggested that the N-terminus (2-100) of RPC4 might interact with the downstream DNA. This region is positioned close to the upstream DNA binding cleft of Pol III and is predicted to bind DNA. These hypotheses can also be tested with the recombinant human Pol III system by systematically deleting the WH domains or the N-terminal part of RPC4 and by testing the effect of these truncations in *in vitro* transcription assays.

The heterodimer expression plasmid is part of a 14-gene pBig2 expression plasmid, which is co-expressed together the heterotrimer genes encoded by a separate pBig1 plasmid (Fig. 2.19 b). To generate the heterodimer truncations variants, I first had to clone the Pol III heterodimer truncations into pBig1e intermediate plasmids shown in Fig. 2.19 c. When I inspected the coding gene of subunit RPC5, I noticed that the original cDNA clone for RPC5, which I used for the cloning, expression and purification of original human Pol III complex, turned out to be a splice variant (Uniprot: Isoform-2; ID-Q9NVU0-2) that lacks the fourth WH domain. I, therefore, inserted the missing residues (649-690) into the originally used RPC5 gene via PCR to create the RPC5 Isoform-1 gene, which I used for cloning of the pBig1e plasmid RPC5-WT-(Isoform-1), RPC4, RPC10 (P14). The RPC5 full-length gene was also used to generate the truncation variant of RPC4 (P15), which I designed to test if RPC4 binds the downstream DNA in context of the human Pol III PIC. To test if the WH domains of RPC5 interact with the SNAP complex and BRF2, I designed RPC5 variants (P16-P19) that also harbour an N-terminal Strep-II tag (Fig. 2.19 c). Thereby, anti-Strep affinity purification should guarantee that no endogenous full-length RPC5 is co-purified during the purification procedure. All cloned heterodimer intermediate pBig1e plasmid show the expected SwaI digestion pattern (Fig. 2.19 d). I also confirmed by Sanger sequencing that the genes were assembled correctly and that the heterodimer variants contained the truncations of interest.

Testing the effect of the heterodimer truncations *in vitro* not only requires cloning of the pBig1e intermediate plasmids into a 14-gene pBig2 expression plasmid via biGBac and co-expressing it with the recombinant heterotrimer (Fig. 2.20 a,b), but also cloning and expression of the 5-subunit SNAPc (Fig. 2.20 c) and of the human TFIIIB subunits TBP, BRF2 and Bdp1 (Fig. 2.20 d). I designed the human SNAPc and TFIIIB expression constructs according to a published biochemical study of the human SNAPc [65].

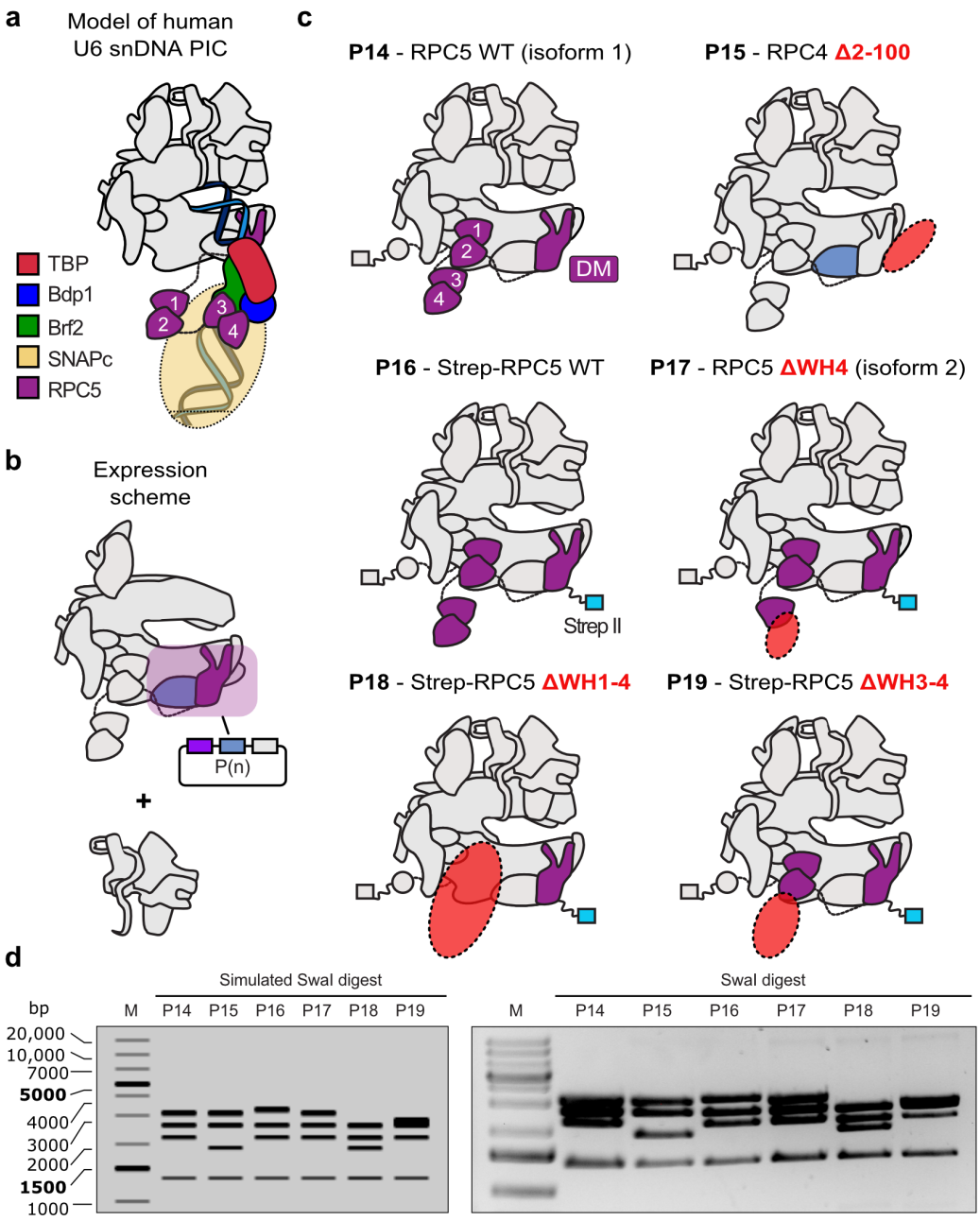

Fig. 2.19. Human Pol III heterodimer truncation constructs. **(a)** Schematic of my proposed model of the human Pol III U6 snDNA pre-initiation complex (PIC). **(b)** Planned expression scheme of human Pol III heterodimer variants. Heterodimer constructs are part of a larger pBig2 expression plasmid (Pol III core + heteodimer) but need to be first cloned into pBig1e intermediate constructs. **(c)** Schematics of human Pol III variants expected to be expressed in HEK293 cells. Plasmid names correspond to pBig1e heterodimer truncations constructs. **(d)** Simulated (left) and real (right) agarose gels (1%) showing SwaI digestion of human Pol III heterodimer truncation constructs P14-P19. M - Marker. Gels were cropped for clarity.

Accordingly, I designed the full-length SNAPc expression construct for co-expression in insect cells with SNAP3 harbouring an N-terminal 6xHis-tag and SNAP4 containing an N-terminal Strep-II-tag (P20). The TFIIIB components were previously expressed in *E. coli* [65]. Of the three TFIIIB subunits, two were already available in the plasmid database of the Müller lab (pET24-6xHis-TBP) and (pET30a-6xHis-Bdp1(1–470)), and I cloned the human BRF2 gene (ordered as cDNA clone) with a C-terminal 6xHis tag into the pET24a vector (P21) for expression in *E. coli*.

To clone the human Pol III heterodimer variants into the pBig2abcde expression vectors, I assembled the pBig1 plasmids P1, P2, P3 and an empty pBig1d vector with one of the heterodimer variants P14-P19 (see Fig. 2.19 c) via biGBac. Of these six assembly reactions, I could obtain four positive clones (illustrated in Fig. 2.19 b) that showed the expected SwaI digestion patterns (Fig. 2.19 e). P20 resembles the human Pol III WT core + plus the WT RPC4-RPC5 heterodimer harbouring all four WH domains. Construct P21 lacks the N-terminus of RPC4. In P22, all WH domains are truncated, and P23 lacks the WH3 and WH4 domains. For the cloning of the SNAPc, I first assembled SNAP1, SNAP2, 6xHis-SNAP3 and SNAP5 into a pBig1a plasmid and the largest SNAPc subunit, Strep-SNAP4, into pBig1b. Second, I joined the two pBig1 plasmids into the pBig2ab SNAPc full-length expression construct, P24 (Fig. 2.20 c), which also showed the expected SwaI digestion pattern. Of all described constructs, the assembly of the pBig1 vectors and of the pBig2 Pol III WT expression plasmid P20 was validated via Sanger sequencing. For two heterodimer variants, P16 and P17, I could not obtain positive clones. Because the questions raised above can, however, be also answered without these constructs, I didn't pursue with further attempts to clone the pBig2 plasmids for these constructs.

In summary, I successfully cloned 24 multi-gene expression constructs for the recombinant production of WT or mutant human Pol III complexes and its multi-subunit co-factor SNAPc. For time reasons, I could, so far, only express and purify the WT Pol III with the RPC5 isoform that lacks the last WH domain (see previous chapter). Hence, the newly cloned Pol III variants and the human Pol III co-factors first need to be expressed and purified. Still, the successful assembly of these truncation or mutant Pol III variants, is the first important step towards a systematic biochemical analysis of human Pol III transcription. Given that the purification of WT human Pol III (RPC5 isoform 2), after some optimization steps, worked fairly well and that protocols for the purification of SNAPc and the human TFIIIB are available [65], I expect that I can obtain these proteins in a timely manner. Hence, I assume that I will soon be able to perform an in-depth testing of the proposed hypotheses that I have outlined above.

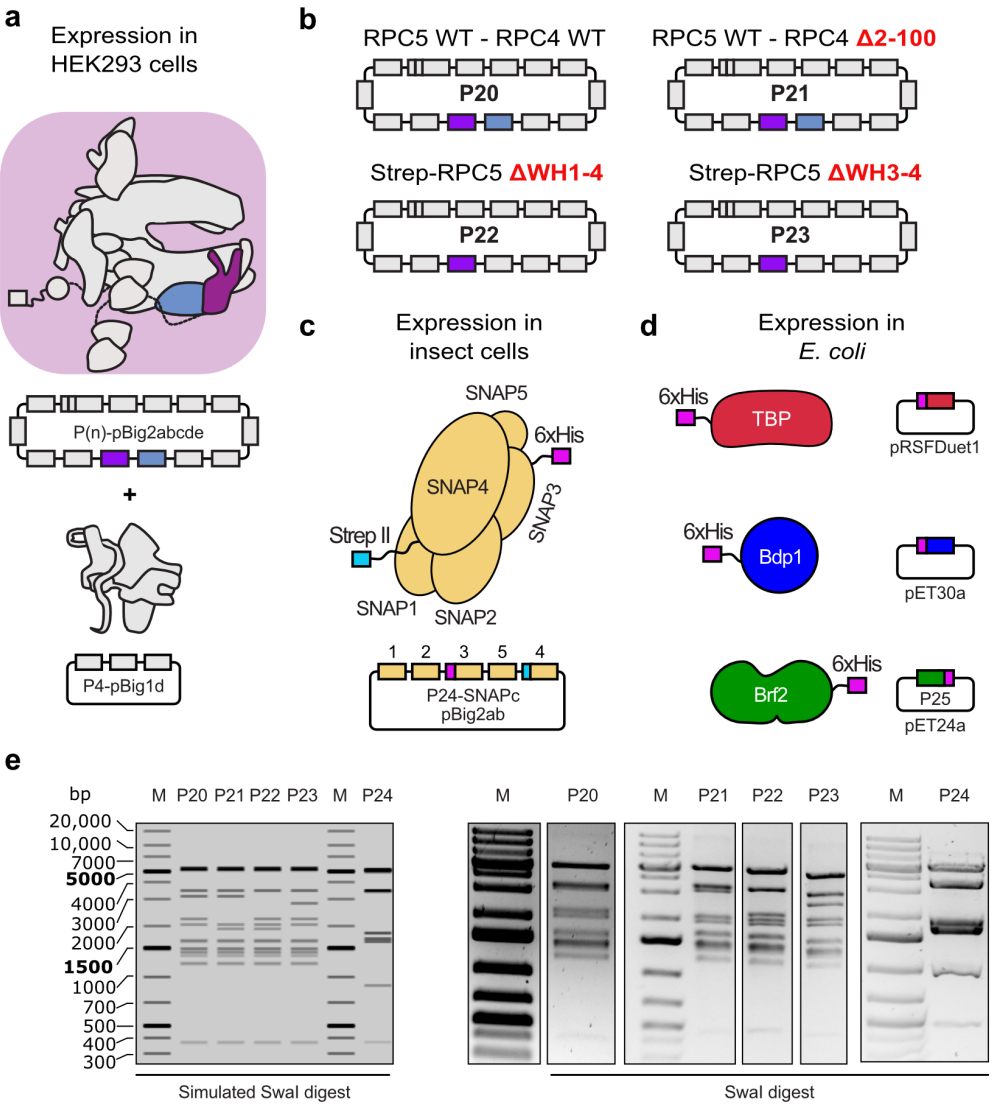

Fig. 2.20. Expression constructs of full-length human Pol III constructs and its co-factors (a) Expression scheme for the production of full-length human Pol III in HEK293 cells. The heterodimer variants listed in Fig. 2.19 c were used to clone the larger P20-P23 pBig2 variants shown in **(b)**. **(c)** Schematic of the SNAPc (top) and its cloned pBig2 plasmid P24 for protein production in insect cells. **(d)** Schematics of the human TFIIIB components TBP, Bdp1 ((1–470), BRF2 (P25) for protein production in *E. coli*. **(e)** Simulated (left) and real (right) agarose gels (1%) showing SwaI digestion of human Pol III pBig2 expression constructs P20-P23 and of the full-length SNAPc. M - Marker. Gels were cropped for clarity.

2.3. Structural and functional characterization of human Pol III - conclusions and outlook

In this section, I have presented a detailed structural analysis of human Pol III in different conformations. The overall architectures of yeast and human Pol III are conserved but the cryo-EM structures of human Pol III shed light on several species-specific differences: (i) a different conformation of the RPC10 (C11 in yeast) CTD; (ii) extended structures of human RPC5 and of the DM domain of RPC4; (iii) the FeS-binding domain of RPC6; and (iv) a potential regulatory role of RPC7α.

Because the reported cryo-EM structures of human Pol III are, resolution-wise, of higher quality than previously reported Pol III structures from yeast [16, 77, 80, 100, 126] and because some elements or conformations appear to be more stable in human Pol III, the structures of human Pol III also contribute to a better overall understanding of Pol III transcription. For example, the human Pol III 'RPC10 inside funnel' conformation revealed how the CTD of RPC10 inserts into the active site to place the acidic hairpin, responsible for the hydrolytic RNA cleavage, close to the 3'-end of the RNA. RNA cleavage also occurs in yeast where it is performed by C11 (the yeast ortholog to human RPC10) and is linked to transcription termination [69, 85]. The structure of the human Pol III 'RPC10 inside funnel' conformation, therefore, gives general insights into Pol III transcription termination as described in the next section in more detail.

Another element that is better resolved in human Pol III and, most likely, has an universal role in Pol III transcription is the stalk bridge, formed by RPC7 (C31 in yeast). Subunit RPC7 is part of the heterotrimer subcomplex. A central region of RPC7 (73-99) protrudes out of the heterotrimer to contact the stalk and folds back to engage with the other heterotrimer subunits again. In the reported yeast Pol III 3D reconstructions, the stalk bridge is also visible at lower threshold levels but the quality was not good enough to assign the register of the stalk bridge with high confidence. The human Pol III structures showed that RPC7 engages with the stalk via a cation-π interaction that appears to be stable because this interaction is maintained upon the conformational rearrangement between apo and elongating Pol III although stalk and heterotrimer move into opposing directions. This observation sheds light on an universal role of RPC7 in coordinating the movements of the stalk and heterotrimer domains and, thereby, potentially in Pol III transcription initiation. In the heterotrimer, RPC7 tightly engages with the WH1, WH2, and CC domains of RPC3 and with the FeS of RPC6. This finding suggests that the primary function of these domains is to anchor RPC7 to the heterotrimer to facilitate such coordinating role of RPC7. This observation also gives insights into the evolutionary origin of the Pol III heterotrimer. The WH1 domain of RPC3 is homologous to the

WH1 domains of the Pol II general transcription TFIIEα and archaeal TFEα [130]. In a related manner, subunit RPC6 contains WH domains that are homologous to the Pol II TFIIEβ protein and archaeal TFEβ, with which it also shares the homologous FeS domain [130]. In contrast, RPC7 has no apparent homologs in archaea and, therefore, most likely was an eukaryotic invention. This suggests that eukaryotes have evolved the Pol III-specific heterotrimer in such way that its RPC7 subunit is capable of coordinating the conformational changes of its mobile elements, namely the stalk, the heterotrimer, and the clamp domain of the Pol III core (which interacts with heterotrimer). To ensure that RPC7 is tightly connected to the core of the heterotrimer, conserved elements of RPC6 (the FeS domain) and of RPC3 (CC, WH1, and WH2) were utilized. Whereas the WH1 domain of RPC1 was already present in archaeal TFEα, the WH2 domain RPC3 presumably evolved via gene-internal duplication from the same origin [81], and I speculate this event occurred specifically for the sake of anchoring RPC7 more strongly to the heterotrimer.

Several disease-associated mutations have been reported for the human Pol III subunits RPC1, RPC2, RPC3, RPAC1, RPAC2, RPC5, and RPC10. The cryo-EM structures of human Pol III allowed mapping these mutations to better understand on a molecular level how these mutations affect Pol III functioning [162] (Fig. 2.21). Agata D. Misiaszek has performed a comprehensive analysis of the disease-associated mutations that is described in more detail in [162]. In brief, to cite the work conducted by Agata D. Misiaszek, the majority of mutated residues that are associated with HLD, WDRTS and TCS lie in the core of Pol III. As such, most of these residues affect the assembly or stability of the Pol III core. In contrast, mutations associated with susceptibility to severe infections by VZV cluster on the surface of Pol III and, potentially, interfere with the sensing of viral DNA in the cytoplasm by disrupting binding sites with co-factors. To similar conclusions came other recently reported studies on the structural characterization of human Pol III [180, 181].

While the cryo-EM analysis was done on endogenous human Pol III, I also discussed that the recombinant human Pol III system can be a powerful tool to functionally characterize transcription by human Pol III. Mutagenesis of human Pol III subunits of interest allows to experimentally test many of the hypotheses that I proposed based on the structure of human Pol III. Furthermore, the recombinant Pol III system could also be used to experimentally analyse the effects of disease-associated mutations on the stability and catalytic activity of human Pol III.

Finally, the structure of human Pol III will serve as a starting point to further explore human Pol III as a drug target. Tumor cells show increased expression levels of human

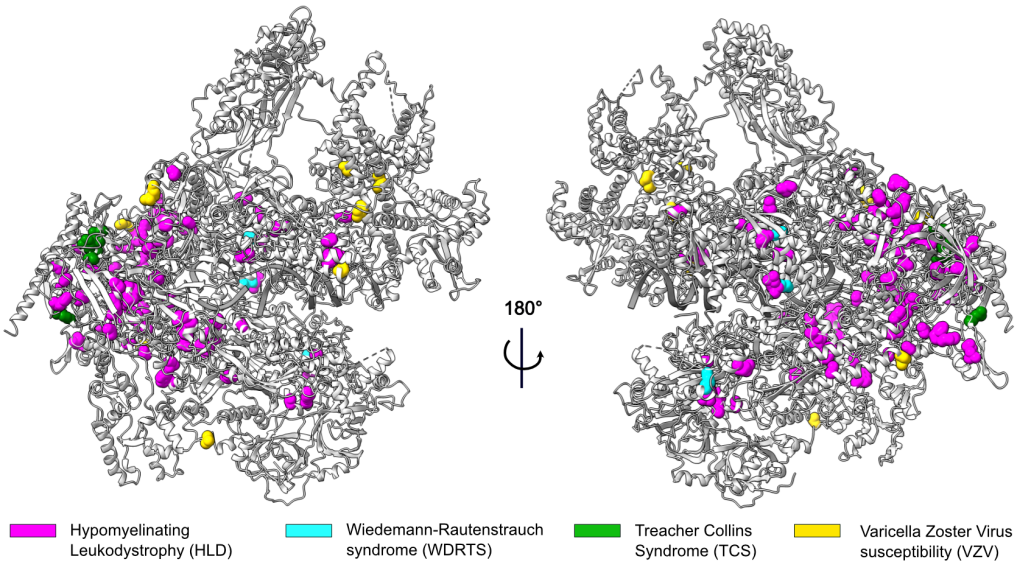

Fig. 2.21. Disease-associated mutations of human Pol III. Structure of human Pol III with mapped disease-associated mutations shown as spheres and colored as follows: pink - Hypomyelinating leukodystrophy (HLD), turquoise - Wiedemann–Rautenstrauch syndrome (WDRTS); green - Treacher Collins syndrome (TCS); yellow - Varicella zoster virus (VZV). The figure has been prepared by Agata D. Misiaszek and is taken from [162].

Pol III transcripts and its transcription factor TFIIIB [111] and the targeting of human Pol III suppressed the growth and viability of cancer cells [133], which renders Pol III a potential drug target to combat cancer. The controlled inhibition of Pol III may also be beneficial for the life expectancies of humans because inhibiting Pol III has been shown to increase longevity in various animals [115]. The central role of Pol III in the eukaryotic cell and the close homology to its paralogous structures Pol I and Pol II, however, requires an in-depth investigation of the compound-assisted inhibition of human Pol III. For this purpose, the high-resolution structures of human Pol III will be highly beneficial because they pave the way towards structure-based design and optimization of lead compounds to specifically target human Pol III. A potential starting point would be to structurally characterize the complex of human Pol III bound by the compound ML-60218. This compound has been identified ca. 18 years ago [182] to be a Pol III-specific inhibitor and, since then, has been widely used to characterize the effects of human Pol III inhibition (e.g. see [133, 183]. However, this compound only has a moderate half maximal inhibitory concentration (IC_{50}) of 27 μM and the structural basis of the mode of inhibition is not known. Hence, a structure of human Pol III bound by ML-60218 could enable to engineer this compound further so that it can bind and inhibit Pol III more strongly while remaining specific for Pol III.

3. Pol III transcription termination

3.1. *In vitro* analysis of Pol III transcription termination and pausing

In this section, I present the structural insights into Pol III transcription termination, which I obtained via cryo-EM. Pol III transcription termination includes the formation of the pre-termination complex (PTC) upon recognition of the 4th dT [92] but the molecular basis of how PTC formation triggers Pol III termination is not known. Therefore, I first aimed to solve the structure of the yeast Pol III PTC.

To ensure that yeast Pol III, purified according to the lab's standard procedure, specifically terminates transcription on poly-dT stretches, I first designed a promoter-dependent *in vitro* transcription termination assay. I used a fusion construct consisting of the *S. cerevisiae* U6 snRNA promoter and the 3'-flanking region of the *S. cerevisiae* tRNA gene tL(CAA)L, which harbours a strong 5-dT termination signal [184] (Fig. 3.1 a). In the presence of yeast TFIIIB and NTPs, Pol III synthesises RNA in a time-dependent manner as shown by the appearance of a prominent double band (Fig. 3.1 b). The size of the RNA double band approximately conforms with the expected size of 32 nucleotides if Pol III specifically terminates on the 5-dT termination signal. To test if the observed RNA corresponds to a specific termination product, I next varied the number of dTs on the NT-strand (Fig. 3.1 c). Changing the 5-dT termination signal to 3- or 4-dTs strongly reduced the RNA signal, observed for 5-dTs and, instead, yielded longer termination read-through products at a size of ca. 50 bases. These results indicate that yeast Pol III, purified according to the lab's standard procedure, indeed terminates on poly-dTs termination signals.

Next, I used RNA primer extension assays to identify a transcription scaffold that traps Pol III in the PTC state. The transcription scaffold consisted of a T-strand, a NT-strand and an RNA primer, which anneals to the T-strand (Fig. 3.2 a). The T-strand was mutated to introduce a 12 nucleotide mismatch, whereby which the melted transcription bubble in the active centre of the polymerase could be mimicked. The RNA 3'-end in the active centre marks the position i-1. Upon addition of NTPs, Pol III extends the annealed RNA, which can be visualized by radioactive labeling of the RNA primer (Fig. 3.2 b). By varying the number of dT-bases on the NT-strand, I identified a 7-dT scaffold (termed SC1) that showed a strong band at the position i+3 upon addition of NTPs, which was absent on a 2-dT EC scaffold. This observation suggested that the poly-dT stretch pauses the RNA-synthesis activity of Pol III and induces PTC formation. To confirm that the number and position of the dT-bases are critical for pausing, I tested another set of transcription scaffolds with a base-by-base variation of the dT-stretch

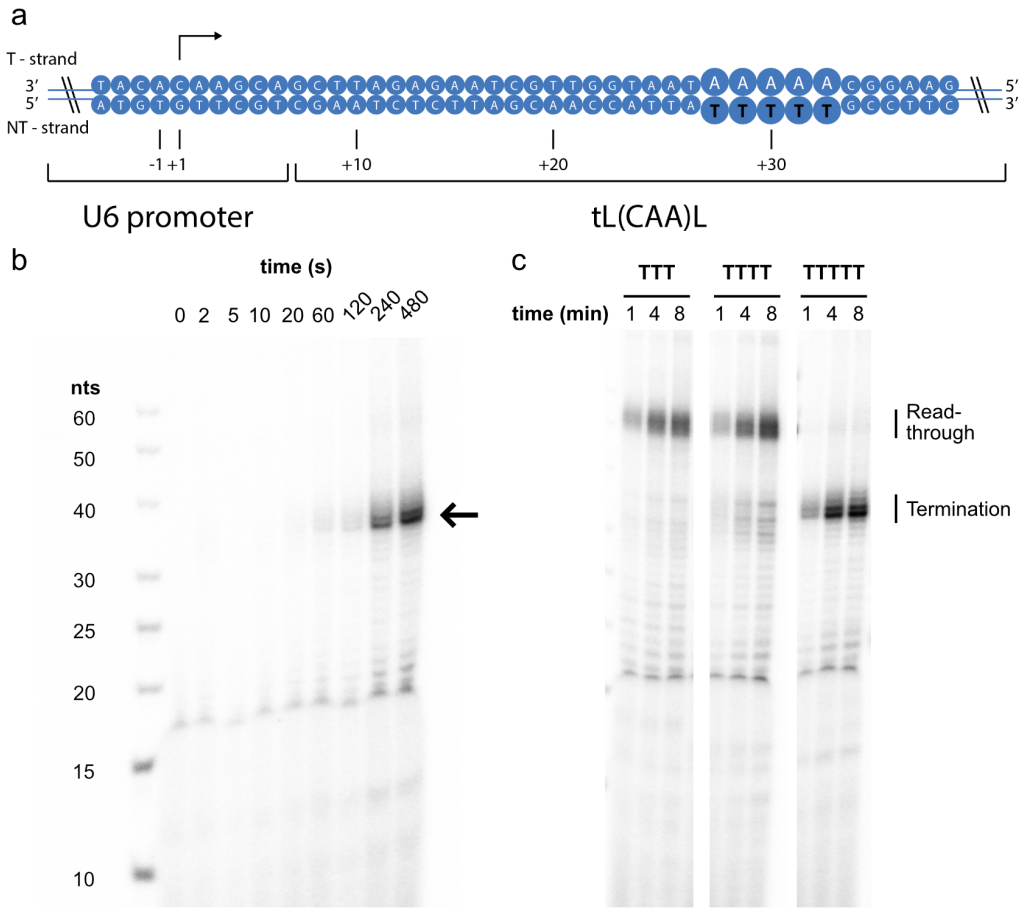

Fig. 3.1. *In vitro* analysis of Pol III transcription termination. (a) *S. cerevisiae* U6-tL(CAA)L fusion construct used for *in vitro* transcription termination experiments. The 5-dT termination signal is highlighted in bold. **(b)** Time course of Pol III transcription using 2.5 pmol U6-tL(CAA)L, 5 pmol TFIIIB and 5 pmol Pol III. The black arrow highlights the radioactive RNA product roughly corresponding to the expected size of specific termination products. Left lane: DNA-ladder; nts - nucleotides. 15% TBE denaturing PAGE. **(c)** Time course of Pol III transcription using a termination sequence of 3 dT, 4 dT, and 5 dTs. Specific termination and transcription read-through products are labeled. 15% TBE denaturing PAGE. The gel has been cropped for clarity.

(Fig. 3.2 c). No stalling was observed on SC3-SC5 that contained less than 5 dTs. In contrast, SC6-SC8 which contained up to 6 dTs in a row showed strong pausing bands, whereas interrupting the poly-dT stretch with a single non-dT base (SC9-SC11) released the paused state again. This results confirmed that the length and the position of the dT-stretch on the NT-strand is critical for Pol III pausing and that the SC1 scaffold likely induces PTC formation.

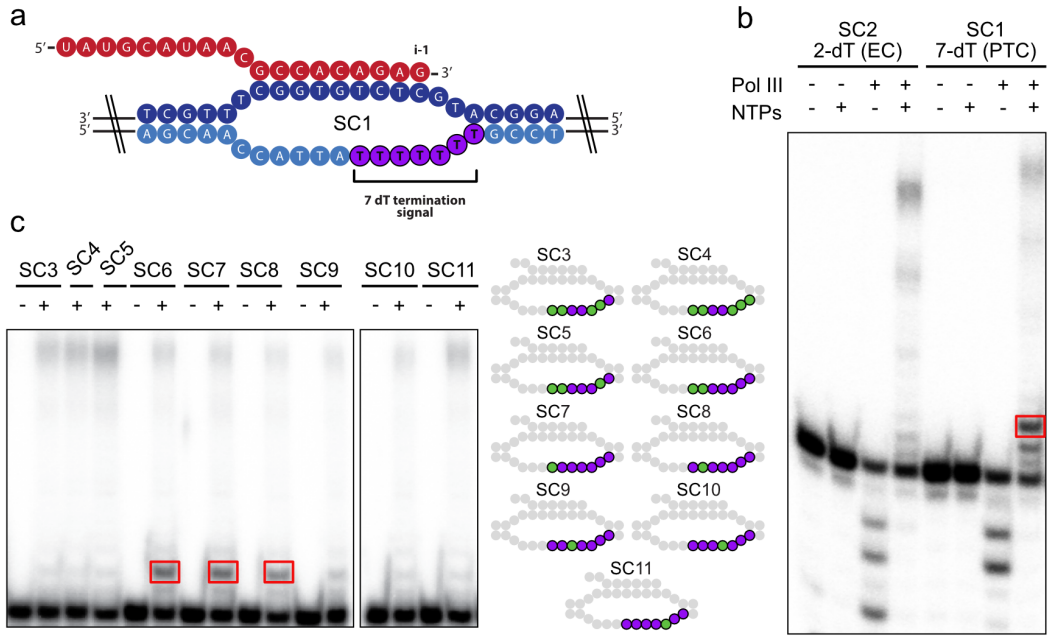

Fig. 3.2. Poly-dT induced pausing of Pol III indicates PTC formation. (a) Schematic of the DNA-RNA scaffold SC1, used in primer extension experiments and for cryo-EM sample preparation of the Pol III PTC. Red: RNA; dark blue: T-strand; light blue: NT-strand; purple: poly-dT stretch. The i-1 indicates the 3'-end of the nascent RNA. **(b)** RNA primer extension assay of a 2-dT elongation complex (EC) scaffold SC2 and the 7-dT PTC scaffold SC1. The PTC scaffold shows a strong band (red frame) at position i+3, which is absent on the EC scaffold. n (Scaffold) = 2 pmol, n (Pol III) = 4 pmol. 15% TBE denaturing PAGE. **(c)** Left: RNA primer extension assay with varying number of dTs. n (Scaffold) = 2 pmol, n (Pol III) = 4 pmol. 17% TBE denaturing PAGE. The red frame labels the i+3 bands also observed on SC1. Right: schematic of analysed scaffolds SC3-SC11. Purple: dT-bases; green: non-dT bases. Gels were cropped for clarity.

3.2. Cryo-EM structure of the Pol III PTC

To solve the structure of the Pol III PTC, I incubated yeast Pol III with SC1 and prepared cryo-EM grids. After I found suitable conditions, which included the supplementation of the sample with 4 mM CHAPSO (see Table 4), I collected a high-resolution dataset of the Pol III PTC on a Titan Krios (Fig. 3.3 a). Processing the data in RELION yielded a 2.8 Å 3D reconstruction (Fig. 3.3 a,b) that locally reached 2.5 Å in the core (Fig. 3.3 c).

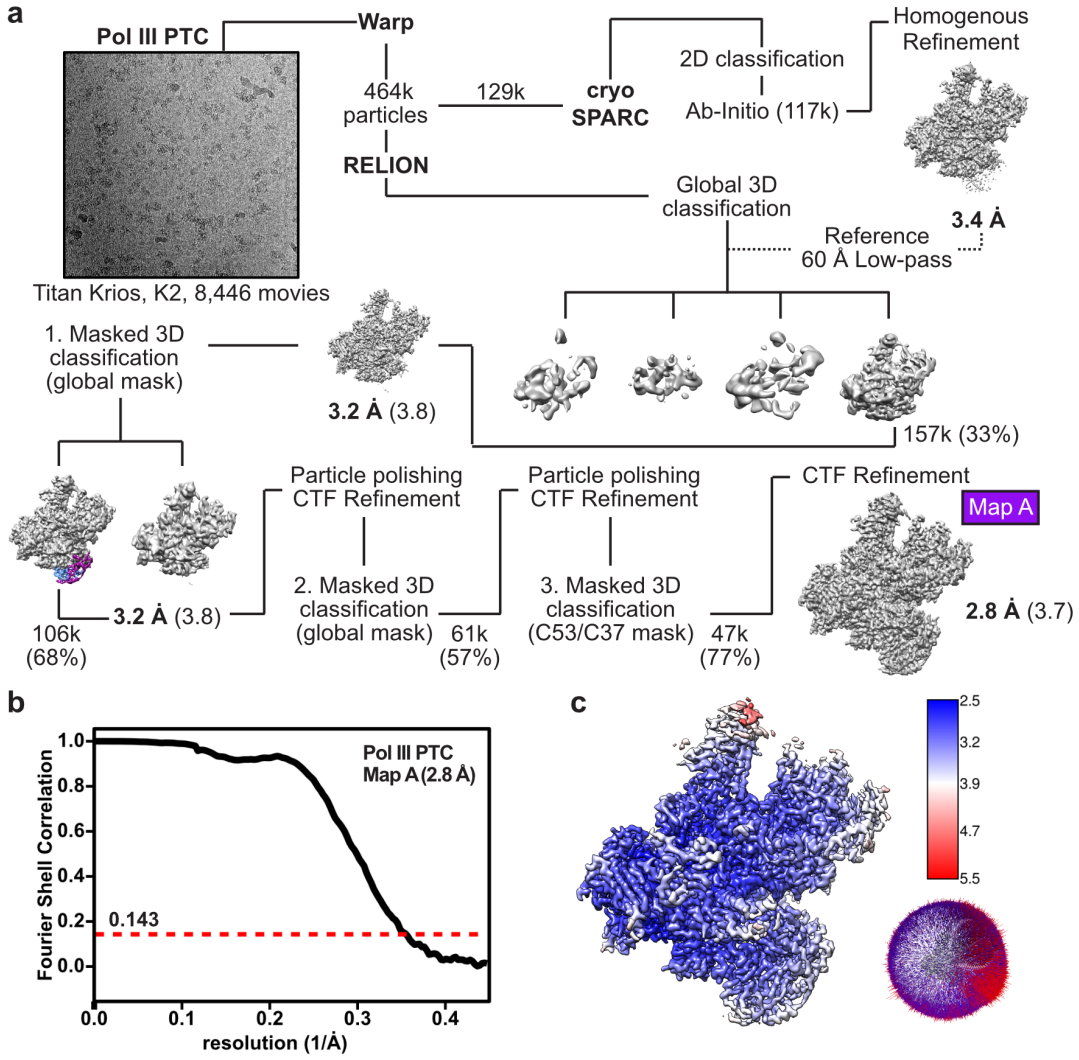

Fig. 3.3. Cryo-EM structure determination of yeast Pol III PTC. (a) Data collection and processing strategy. Resolution values of the sharpened maps are given below and of unsharpened maps in parenthesis. Particle numbers were rounded. **(b)** FSC curve of yeast Pol III PTC (FSC = 0.143) **(c)** Local resolution estimation and angular distribution plot.

In a similar manner, I collected a dataset for the yeast Pol III EC as control, which I prepared under the same conditions, to ensure that any structural rearrangements between Pol III EC and PTC are not due to differences in cryo-EM sample preparations. For the yeast Pol III EC, I used an transcription scaffold that harboured only 2 dTs on the NT-strand, and the obtained 3D reconstruction had a nominal resolution of 3.4 Å (Fig. 3.4 a,b) and reached 2.9 Å in the core (Fig. 3.4 c).

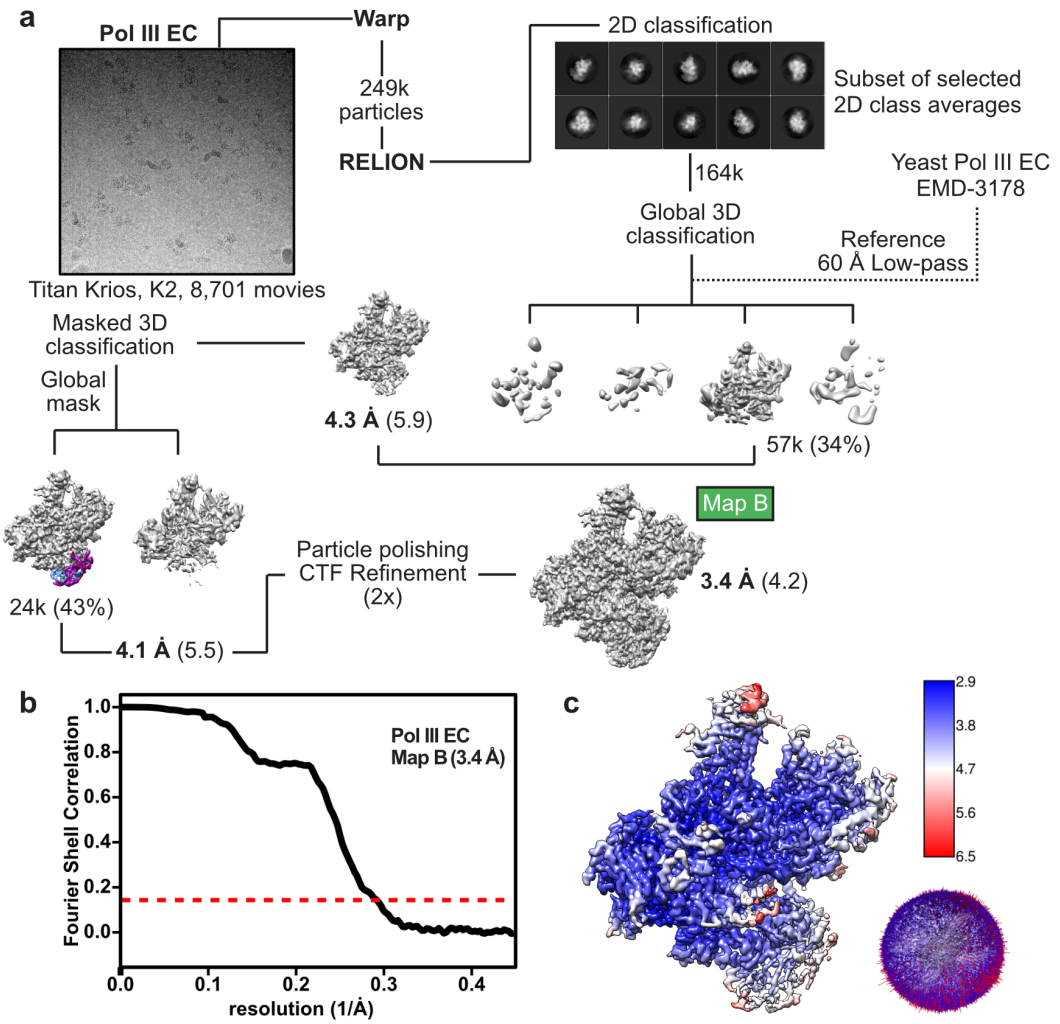

Fig. 3.4. Cryo-EM structure determination of yeast Pol III EC. (a) Data collection and processing strategy is outlined in a similar manner as described in Fig. 3.3. **(b)** FSC curve of yeast Pol III EC (FSC = 0.143). **(c)** Local resolution estimation and angular distribution plot.

The Pol III PTC 3D reconstruction showed well resolved signals for the DNA and the RNA (Fig. 3.5 a,b). Importantly, I observed a strong signal for the DNA NT-strand, which enabled me to build all 7 dTs of the termination signal (Fig. 3.5 c). In contrast, the NT-strand is more flexible and, therefore, mostly not visible or less well resolved in

the EC structures of Pol I, Pol II and Pol III EC control. Not only does the NT-strand become ordered during the transition from Pol III EC to PTC, but it also induces a local contraction of subunit C128, which tightly associates with the NT-strand (Fig. 3.5 d).

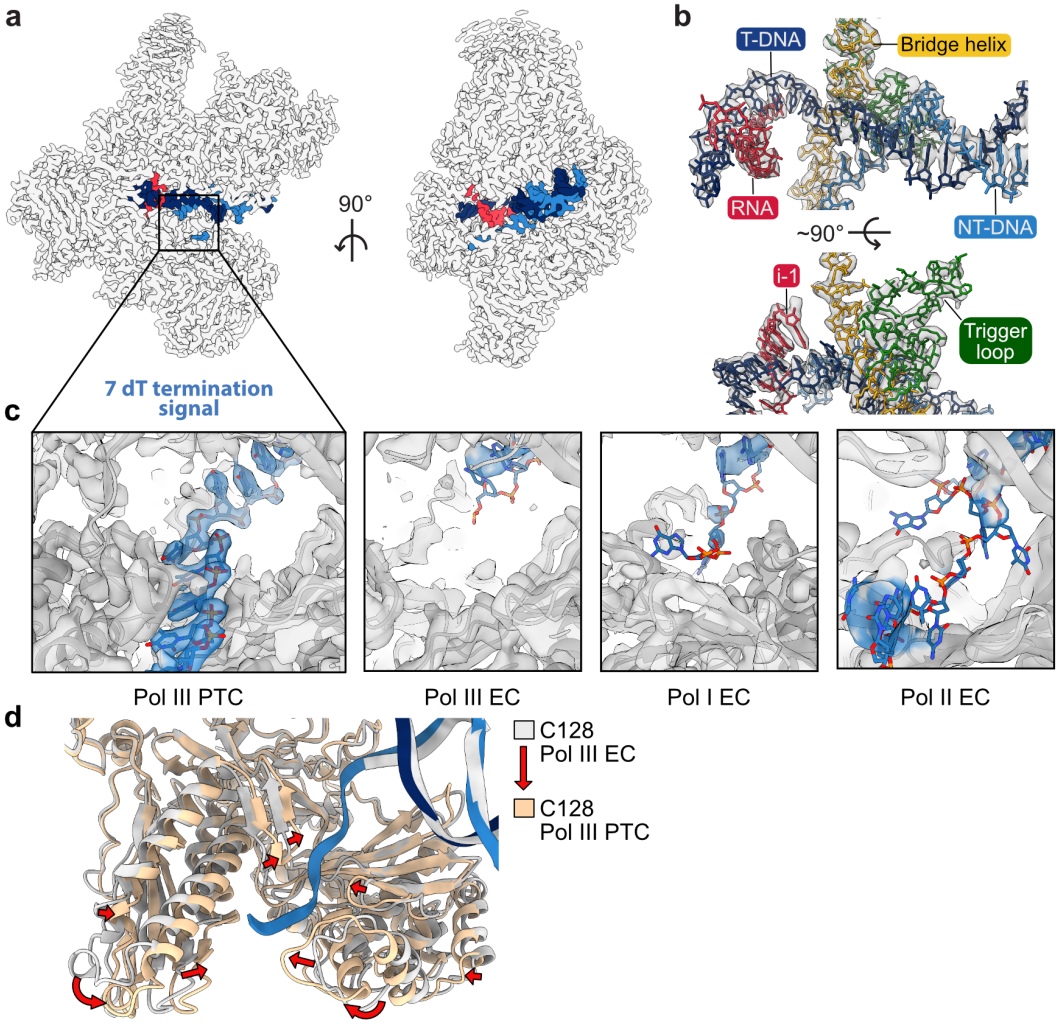

Fig. 3.5. Cryo-EM structure of the yeast Pol III PTC. (a) Cryo-EM map of the yeast Pol III PTC. The DNA and RNA are coloured (dark blue: T-strand; light blue: NT-strand; red: RNA). **(b)** Close-up view onto the active site showing built structural models for nucleic acids, the C160 bridge helix and trigger loop and their fit into the cryo-EM map (transparent grey surface). The RNA i-1 site is labeled. **(c)** Close-up views onto the NT-strands (blue) of the yeast Pol III PTC and of structures of comparison as labeled below. Coordinates: Pol III EC - coordinates and map this work; Pol I EC - PDB: 6HLR, EMDB: 0240; Pol II EC - PDB: 5C4J and related 2Fobs-Fcalc calculated at 1.2 σ. **(d)** Superimposition of Pol III EC (gray) and PTC (coloured). Red arrows illustrate contraction of C128.

3.3. Molecular basis of NT-strand recognition in the Pol III PTC

The Pol III PTC structure allows a detailed analysis of how Pol III recognized the 7 dTs of the NT-strand. The termination signal is exclusively bound by subunit C128, which forms a positively charged groove that contacts the NT-strand (Fig. 3.6 a). A closer inspection of the contact points reveals that 14 hydrogen bonds (H-bonds) form between C128 side chains and all 7 dTs of the termination signal (Fig. 3.6 b, d). Both the NT-DNA phosphate backbone and dT bases form H-bonds with C128. The H-bond forming residues can be mapped to the C128 lobe (197-370), protrusion (371-439) and fork (440-521). In addition, C128 forms a hydrophobic cavity that accommodates the dT-specific methyl groups of the third and fourth dT, illustrating how Pol III can, potentially, distinguish between a cytosine and a thymine base (Fig. 3.6 c).

To understand the importance of the NT-strand-contacting residues further, I examined the degree of sequence conservation of these amino acids. I performed multiple sequence alignment (MSA) of C128 orthologs from 15 eukaryotic species that cover five eukaryotic supergroups: stramenopiles (2), alveolata (2), amoebozoans (1), archaeplastids (3), ophistokonts (7) (see Fig. A1 in appendix). Next, I subjected the generated MSA to the Scorecons software [185] to retrieve conservation scores of the respective residues. Fig. 3.6 d shows the nucleotide blot of the C128-NT-strand interaction network, in which the H-bond forming residues are colored by the obtained conservation scores. Four residues (Q199, K228, K393, R446) are identical in all supergroups and five residues exhibit a high conservation score of 0.8-0.9 (S229, N245, R451, Q476, R481), suggesting that these residues are particularly important for NT-strand recognition. The residues K230, S246, K448, E450 show a moderate conservation score of 0.5 to 0.7 and only one residue (T311) is not conserved at all.

Next, I performed an MSA of *S. cerevisiae* C128 (Pol III) and its paralogs RPB2 (Pol II) and A135 (Pol I) (see Fig. A2 in appendix) and analysed to what extent the H-bond forming residues are unique to Pol III or, alternatively, are also present in Pol II and Pol I. As shown in Fig. 3.6 e, 9 out of 14 residues are unique to Pol III (Q199, K228, N245, S246, T311, K448, E450, Q476, R481). Four residues can be found in both Pol II and Pol III (K230, K393, R446, R451) and only one residue (S229) is conserved between all three Pols. Hence, to form H-bonds with the poly-dT termination signal on the NT-strand, the Pol III subunit C128 utilizes a set of residues that exhibit high conservation scores (9 out of total 14 residues), of which the majority (6 out of 9 residues) is unique to Pol III.

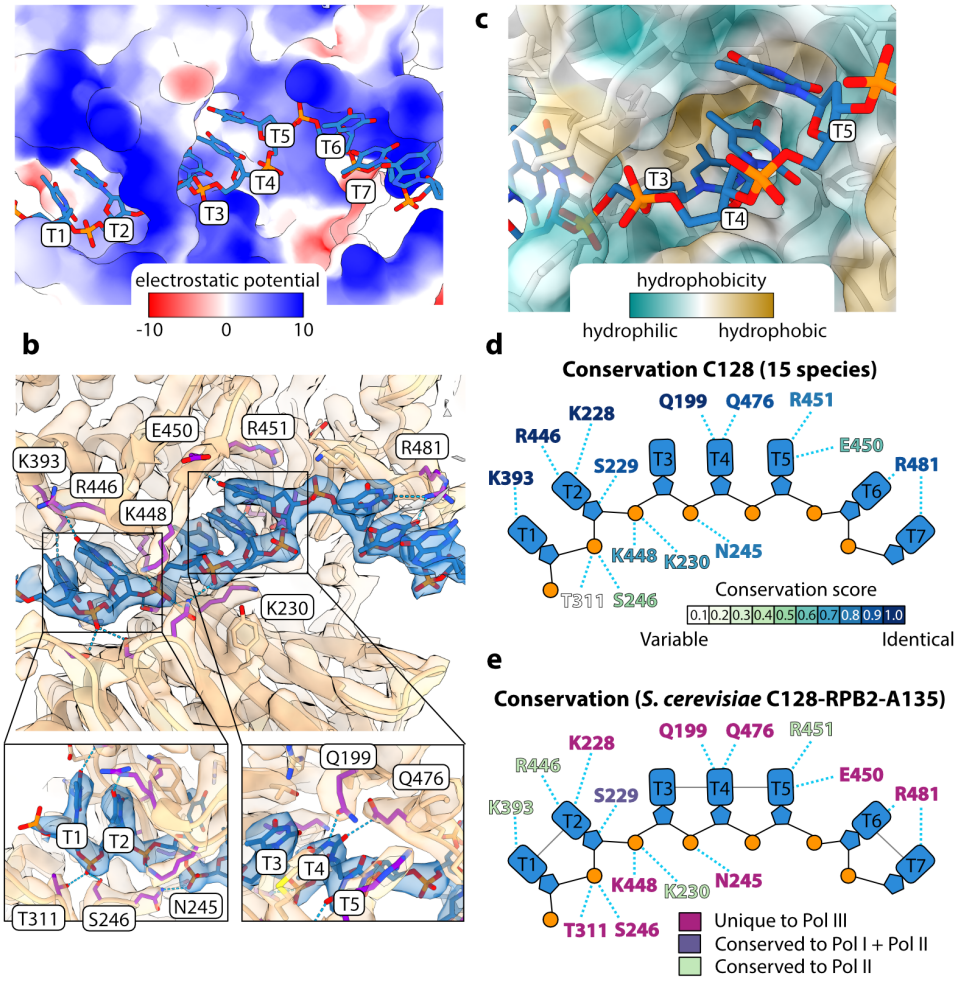

Fig. 3.6. NT-strand recognition by subunit C128 in the Pol III PTC. (a) C128 in Pol III PTC is shown as a surface and colored by electrostatic potential (red: negative, blue: positive). The NT-strand runs along the positively charged groove of C128. **(b)** Hydrogen-bond (H-bond) network formed between C128 and the 7-dT termination signal of the NT-strand. C128 is shown as wheat-colored ribbon and residues forming H-bonds with the NT-strand are shown as sticks and are labeled. H-bonds are shown as blue dotted lines. The cryo-EM map of the Pol III PTC is shown as transparent surface and is coloured in wheat (C128) and blue (NT-strand). **(c)** Forkloop 1 of C128 forms a hydrophobic cavity that accommodates the methyl groups of T3 and T4. The model of C128 is shown as ribbon and its corresponding surface (transparent) is coloured by hydrophobicity. Cyan: hydrophilic, gold: hydrophobic. **(d)** Nucleotide blot of the C128-NT-strand H-bond network shows conservation of H-bonds forming residues across the eukaryotic kingdom. Conservation scores were obtained by performing multiple sequence alignment (MSA) of C128 orthologs from divergent 15 species (see supplement) and subjecting the MSA to the Scorecons software [185] using C (entropy) with 21 types as scoring method. **(e)** Conservation of NT-strand contacting residues of *S. cerevisiae* Pol III C128 and its paralogs in Pol II (RPB2) and in Pol I (A135). Unique residues in Pol III C128 either have gaps in the MSA between the three paralogous sequences or are of chemically different nature. Residues that are conserved across the three paralogs are either identical or of similar chemical nature.

3.4. Fork loop 1 and 2 recognize the NT-strand and define its path

Several key residues for NT-strand recognition and coordination can be mapped to the fork region of C128 (440-521) and, more specifically, to fork loop 1 (440-453) and fork loop 2 (475-486). The key residues that lie in the fork loop 1 are R446, K448, E450, R451. R446 forms an H-bond with the base of dT2, K448 an H-bond with the phosphate of dT3, and an H-bond between the base of dT5 can be formed with the peptide backbones of E450 and R451. The fork loop 1 and fork loop 2 elements can also be found in the C128 paralogs RPB2 (Pol II) and RPA135 (Pol I). Interestingly, fork loop 1 in Pol III adopts a different fold or conformation compared to Pol II and Pol I (Fig. 3.7 a). In Pol I, fork loop 1 (469-483) is larger and adopts an open loop-like fold, which would clash with the NT-strand of the Pol III PTC when superimposing the two structures. In contrast, the fork loop 1 of Pol III folds into a β-hairpin although this becomes not apparent in secondary structure prediction analysis. In Pol II, the fork loop 1 (466-478) is not build as a β-hairpin but its overall shape still resembles a similar fold as the fork loop 1 in Pol III. Notably, the two fork loops adopt different conformations. In Pol III, fork loop 1 adopts a closed conformation and, thereby, forms the H-bonds with NT-strand and the hydrophobic cavity that accommodates the dT methyl groups. In contrast, fork loop 1 in Pol II swings open and points away from the NT-strand. Neither the H-bond network with the NT-strand, nor the hydrophobic cavity could form in its current conformation.

Besides the H-bond-forming residues, other amino acids function in coordinating the NT-strand and, thereby, define the path of the termination signal. A comparison between the Pol III EC and PTC reveals an interlock function of R451 (Fig. 3.7 b). Whereas in the Pol III EC, R451 points towards the coiled-coils of the C128 protrusion, R451 flips by approximately 90 °C during the EC to PTC transition. Thereby, it forms two H-bonds with the main chain carbonyls of S475 and Q476, which are both part of the fork loop 2 element. Not only could this switch of R451 be a determinant for the observed compaction of the Pol III during the formation of the PTC but it also seems to function as an interlock by keeping the fork loop 1 in a defined position that, essentially, forces the NT-strand to follow the observed path. The importance of the interlock fixing the position of fork loop 1 becomes further apparent when locking at residue K448, which is also part of fork loop 1. Together with the neighbouring N245, these two residues function as a molecular gate through which the NT-strand is threaded and the base-stacking between dT2 and dT3 is disrupted (Fig. 3.7 c). Fork loop 2 also harbours three key residues for NT-strand recognition. Q476 forms an H-bond with the dT4 base and R481 forms two H-bonds with the bases of dT6 and dT7 (Fig. 3.7 d). Residue K479 does not

form H-bonds with the NT-strand but it functions as a molecular roadblock that breaks the base stacking between dT5 and dT6 and defines the path of the NT-strand.

In summary, Pol III uses fork loop 1 and 2 to recognize the NT-strand and to define its path. Pol III's fork loop 1 contributes several H-bond forming residues and adopts a different fold and/or conformation compared to Pol I and Pol II. Residue R451 interlocks the position of fork loop 1, which allows formation of the hydrophobic cavity and keeps the K448 gate in position. The pathway of the NT-strand is determined by the K448 gate and the K479 roadblock, which both interrupt NT-DNA base stacking between dT2 and dT3, as well as between dT5 and dT6. Thus, the Pol III PTC formation, forms a tight H-bond network and a binding groove for the NT-strand, which has to follow a defined pathway through the Pol III core. All three elements are likely key determinants for termination-signal induced pausing, which is a prerequisite for Pol III termination.

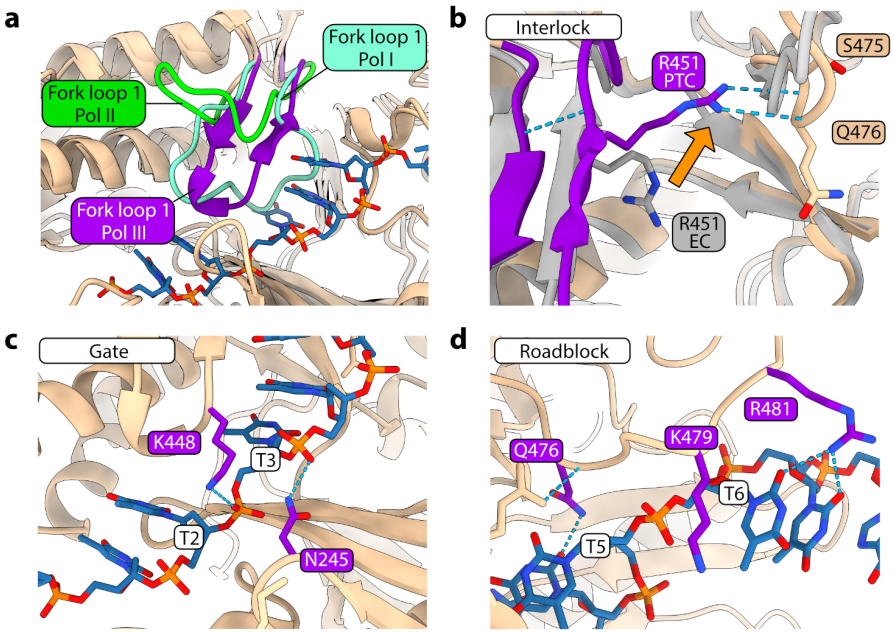

Fig. 3.7. Close-up views on selected elements playing key roles in NT-strand recognition and coordination. (a) Comparison between fork loop 1 of C128 (Pol III), RPB2 (Pol II EC; PDB: 5C4J), and A135 (Pol I EC; PDB: 6HLR). **(b)** Close-up view and superimposition of fork loop 1 of Pol III PTC (purple and wheat) and Pol III EC (grey). Residue R451 changes its conformation during the EC to PTC transition (illustrated as orange arrow) to form two H-bonds with the neighbouring fork loop 2 residues S475, Q476. Thereby, R451 potentially functions as an interlock that holds fork loop 1 in position to tightly engage with the NT-strand. **(c)** Close-up view on Pol III PTC K448 and N245 that form a molecular gate through which the NT-strand is threaded. **(d)** Close-up view on fork loop II and residues Q476, K479, R481 that coordinate the NT-strand. K479 functions as a molecular roadblock that defines the path of the NT-strand.

3.5. Mutagenesis of yeast Pol III for structure-function studies

Based on the Pol III PTC structure and conservation of the key residues in NT-strand recognition, I designed point mutations that are expected to interfere with Pol III pausing and termination. Table 2 lists the selected residues, their functions, the designed mutations and describes the rationale of the chosen substitutions. In all expect one residue (Q476), the MSA of C128, RPB2, A135 was used to replace the C128 residues with their counterparts in RPB2 (Pol II-like) or A135 (Pol I-like). For Q476, the Pol I-like substitution would have been a serine, which could still form an H-bond with T4, and the Pol II-like substitution would be a proline, which might interfere with folding of the fork loop 2. I, therefore, designed the Q476 mutation to be an alanine substitution, which should omit H-bond formation but shouldn't destabilize the fold of fork loop 2. To insert the mutations into the *S. cerevisiae* genome, we used the CRISPR-Cas9 technology, which was performed by Dr. Helga Grötsch (see methods). Notably, 6 out 10 designed mutations were successfully introduced. Two of the failed mutations (Q476 and K479) were still covered by the Triple mutant (Q476A,K479A,R481G) meaning that these residues could still be covered in our structure-function studies. Despite extensive attempts, two designed mutants (K393, R446) did not yield any clones that were positive for the introduced mutations. K393 and R446 form H-bonds with T1 and T2, respectively. The importance of the recognition of the first and second dT bases could, therefore, not be tested further. An explanation for the unsuccessful attempts for K393 might be that no canonical PAM sequence for Cas9 (NGG) could be chosen to target the donor guide to the genomic location. Instead, a non-canonical PAM sequence (NAG) had to be used. For R446, 16 out 16 screened colonies showed silent point mutations, suggesting that the insertion of the intended substitution (arginine to serine) might be lethal.

The mutations were introduced into the *S. cerevisiae* genome of a strain carrying an C-terminal TAP tag on subunit C128, which allowed purifying endogenous Pol III that carried the C128 mutations. I purified both wild-type (WT) and mutant Pol III variants from 16 L yeast cultures, which were grown by Dr. Helga Grötsch. For the purification, I adapted the protocol for large-scale purification of AC40-tagged Pol I and Pol III (100 L, fermenter-cultured) to smaller scale. The protocol included cell lysis in a bead beater, a heparin chromatography step, IgG affinity chromatography and elution from IgG beads with Tobacco Etch Virus (TEV) protease. Because only Pol III, and not Pol I, is affinity-tagged, I omitted the final ion-exchange chromatography step. I could obtain the following yields (in µg) of purified Pol III variants: 48 (WT), 50 (Triple), 111 (Q199R), 12 (K448A), 168 (R451V), 96 (R481G), 87 (H225L). Fig. 3.8 shows the protein gels of the purified Pol III variants. All Pol III mutants show the same band pattern

as Pol III WT (both C128 and AC40-tagged) indicating that the mutants variants are intact and suitable for the structure-function analysis.

Table 2. Key residues of Pol III C128 contributing to termination signal recognition and designed mutations for structure-function studies.

Residue	Function	Mutation	Explanation	Acronym	Success
Q199	H-bond with T4 (base)	R199	Pol I / Pol II-like substitution, charge repulsion, steric clash	Q199R	**Yes**
H225	H-bond with D72, potentially functions in Pol III contraction	L225	Pol II-like substitution	H225L	**Yes**
K393	H-bond with T1 (base)	D393	Pol I-like substitution, charge reversal	K393D	No
R446	H-bond with T2 (base)	S446	Pol I-like substitution, charge removal	R446S	No
K448	Gating function, H-bond with T3 (phosphate)	A448	Pol II-like substitution, removal of charge and atom mass	K448A	**Yes**
R451	Interlocks the position of fork loop 1	V451	Pol I-like substitution, charge removal	R451V	**Yes**
Q476	H-bond with T4 (base)	A476	Charge removal	Q476A	No
K479	Molecular roadblock forcing NT-strand to follow its path	A479	Pol I-like substitution, removal of charge and atom mass	K479A	No
R481	H-bonds with T6 and T7 (bases)	G481	Pol II-like substitution, removal of charge	R481G	**Yes**
Q476, K479, R481	H-bonds with T4, T6, T7 (bases), molecular roadblock	A476, A479, G481	Removal of charge and atom mass	Triple	**Yes**

3.6. *In vitro* and *in vivo* Pol III termination structure-function studies

To test the effect of the Pol III C128 mutants *in vitro*, I designed a tailed-template transcription assay. Pol III can be added to a dsDNA construct that harbours a single-stranded 3'-overhang of the T-DNA strand, which allows monitoring Pol III transcription in the absence of a promoter DNA and transcription factors (Fig. 3.9 a). To monitor transcription termination, I introduced a 6 dT termination signal on the NT-strand. To ensure that observed RNA bands correspond to termination-induced release of the RNA transcript, I also designed a second construct that only contained 3 dTs, through

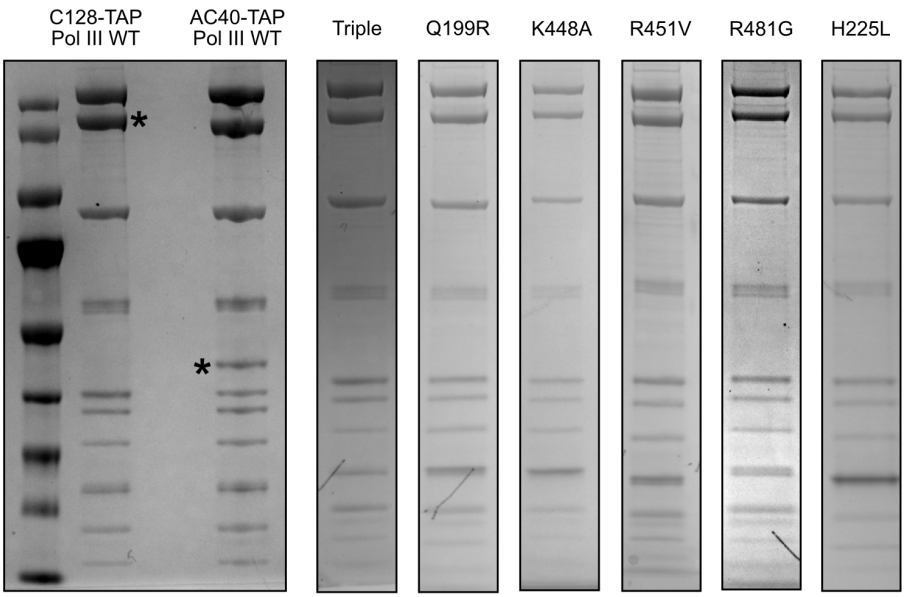

Fig. 3.8. Pol III C128 WT and mutant samples for structure-functions studies. Protein gels of *S. cerevisiae* Pol III WT samples and of C128 mutants. C128-TAP WT and mutants were purified from 16 L cultures. The AC40-TAP version is loaded as a control. The asterisks mark affinity-tagged subunits that still contain the calmodulin-binding domains of the TAP tag.

which Pol III should read through. To precisely know the positions of the termination bands and of the read-through products, I designed the construct to contain only a single cytosine base (dC) downstream of poly-dT termination site. By omitting the use of GTP in the reaction mix, Pol III should accumulate downstream of the termination site at the marked position if it fails to terminate correctly.

Fig. 3.9 b shows the outcome one of the tailed-template transcription assays. I performed the experiments in triplicates, which allowed quantification of the RNA bands (Fig. 3.10). For the WT Pol III, a strong RNA band is observed at the expected position (corresponding to dT5), which is absent in the 3-dT control lane (compare lane 2 and 11), indicating that Pol III WT terminates correctly on the 6-dT terminator. All analysed mutants, except H225L, show a clear reduction of the termination signal (lanes 3-8). The strongest reductions can be observed for Q199R (ca. 4-fold), Triple (ca. 3-fold) and R451V (ca. 3-fold). The variants K448A and R481G also show reduced termination, which is, however more modest (ca. 1.7 fold) and almost no effect can be observed for H225L (ca. 1.2 fold). In addition to a strong reduction of termination bands, the variants Q199R and R451V also show an increase in RT products (ca. 1.7 fold). No clear difference can be observed for the band intensities of the 3-dT control RT products (lanes 11-17). Thus, the observed effects of the Pol III C128 mutants specifically affect Pol III termination

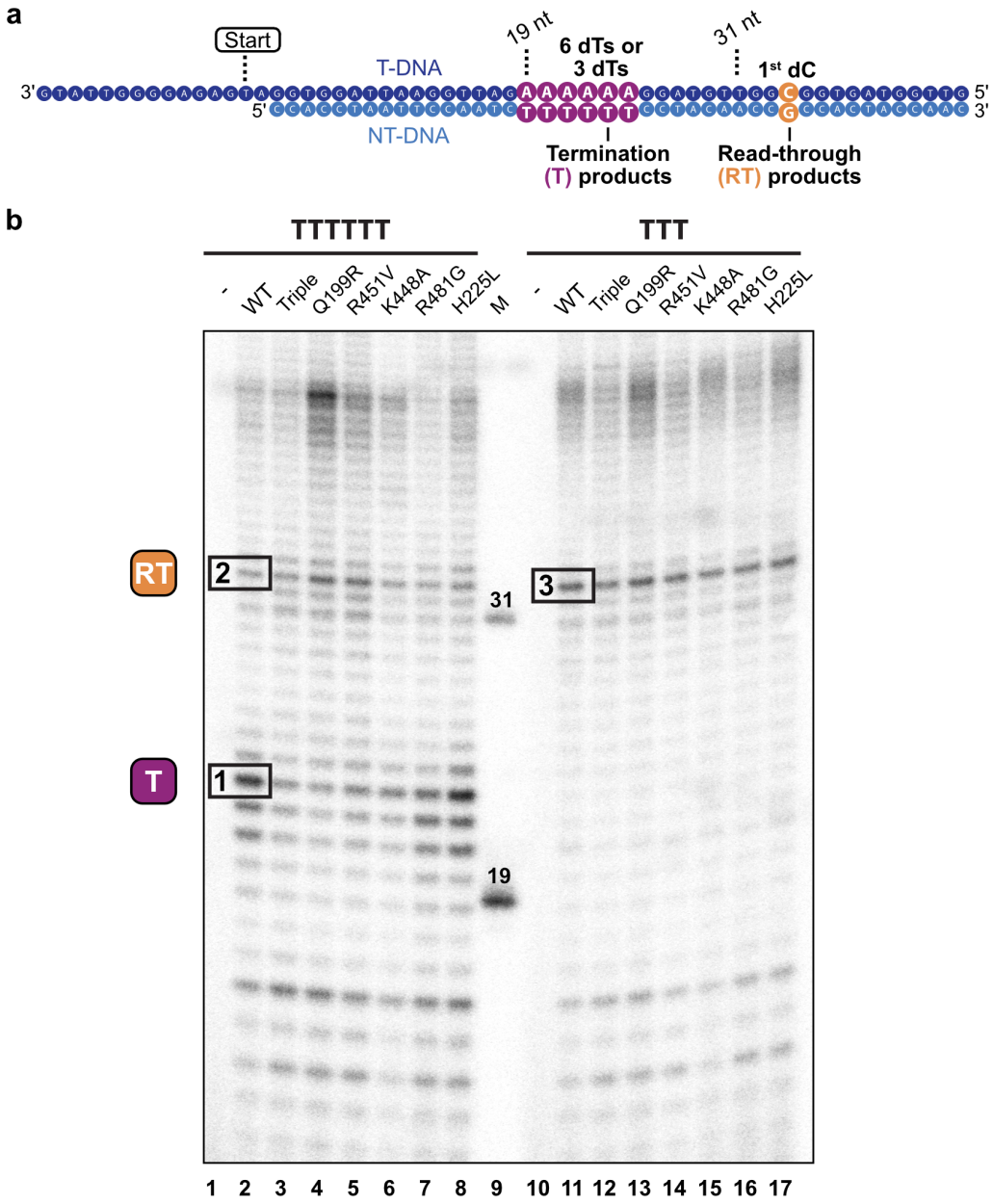

Fig. 3.9. Pol III C128 mutations interfere with Pol III termination *in vitro*. **(a)** Schematic of the used tailed-template DNA construct. A single-stranded T-DNA overhang allows binding of Pol III in the absence of transcription factors. The termination signal is colored in purple and either contains 6 dTs or 3 dTs (as elongation control). The first cytosine (dC) is colored in orange, which should accumulate termination RT products when only using ATP, CTP and UTP in the reaction buffer. Positions of the loaded marker bands (19 and 31 nucleotides (nts)) and the transcription start site are marked. **(b)** Denaturing RNA gel of the tailed-template transcription assay. Observed termination (T) and termination read-through (RT) products are marked. Bands that were used to quantify transcription products are boxed and labeled with 1: T(6 dTs), 2: RT(6 dTs), 3: RT(3 dTs).

but not Pol III elongation. All analysed samples also show levels of full-length products, which may be attributed to a misincorporation of NTPs. Surprisingly, the Q199R mutant shows a strong enrichment for the full-length product (lane 4) suggesting that residue Q199 may also function in transcription proofreading to ensure that only the correct nucleotide is incorporated into the RNA chain.

The tailed-template assays confirm the importance of the residues Q199 (contacts dT4) and R451 (interlocks fork loop 1) to pause Pol III so that it can terminate correctly. The Triple mutant also confirms the importance of residues Q476 (contacts dT4), K479 (roadblock), and R481 (contacts dT6, dT7) for transcription termination but to what extent each of the three residues contributes to the observed reduction of termination cannot be distinguished. Interestingly though, the R481G single point mutant only shows a modest reduction of termination, suggesting that termination may be more affected by the recognition of dT4 (either by Q199 or Q476) or by the roadblock function of K479. The modest effect of K448A implies that K448 also assists transcription termination via its gating function and/or its H-bond with the dT3 phosphate but it appears to be not as crucial as Q199 and R451. Lastly, the H225L mutant was tested because H225 forms an H-bond with the neighbouring Pol III protrusion element, which contracts upon the EC to PTC transition. The fact that the H225L mutant does not show a clear effect on transcription termination suggests that H225 either is not critical for the observed contraction of that the contraction of the Pol III protrusion is rather a consequence of NT-strand recognition instead of a requirement.

To test the effect on the Pol III C128 mutations *in vivo*, Dr. Helga Grötsch performed yeast spotting assays (Fig. 3.11). The Triple mutant does not affect yeast growth at permissive temperatures but, strikingly, shows a severe heat-sensitive growth phenotype with no single colony being detectable at 37 °C. R481G also shows a heat-sensitive phenotype, which is, however, less severe than the Triple mutant phenotype. In contrast, the Q199R mutant already displays a growth defect at 30 °C, which is even more pronounced at 25 °C. The K448A and R451V mutants do not show a growth phenotype.

In summary, the *in vitro* structure-function studies confirm that the NT-strand recognition via key residues in C128 is critical for Pol III transcription termination. Three out of five tested mutations (Triple, Q199R, R481G) also show a growth deficiency in yeast spotting assays, suggesting that interference with Pol III termination also affects yeast cell viability. K448A and R451V don't show any growth defects, which needs to be explored further. Potentially, for some mutants,the effects on transcription termination are buffered by other factors such as the RNA helicase Sen1, which has recently been suggested to function in Pol III transcription termination [150]. We are currently collaborating with

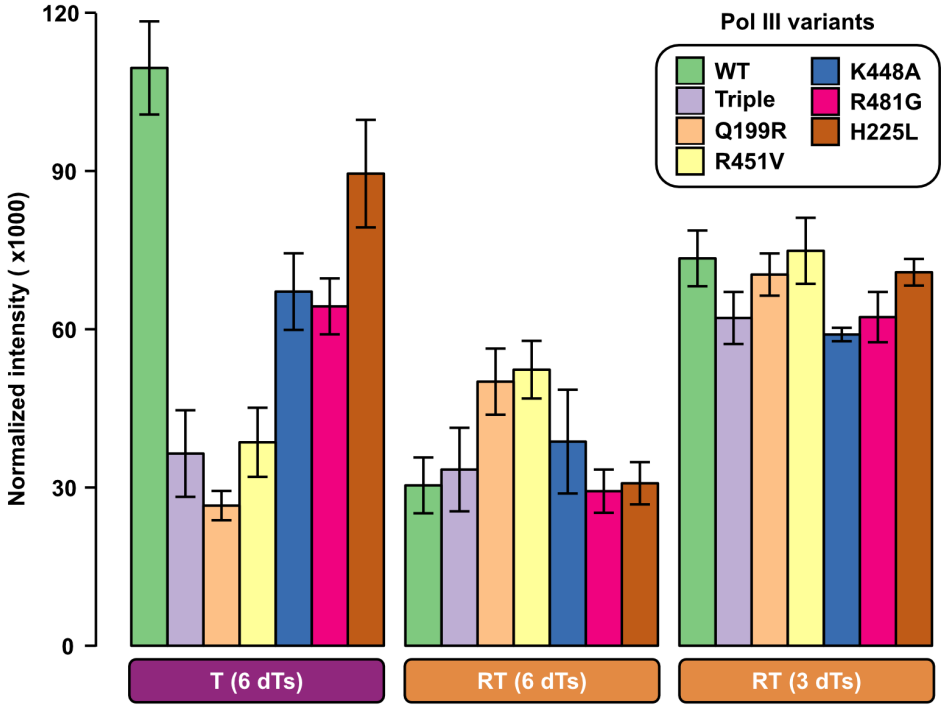

Fig. 3.10. Quantification of tailed-template transcription assays to monitor Pol III termination. Barplot showing normalized intensities of selected RNA bands for the analysed variants (mean ± SD (n=3). RNA band (marked for WT Pol III in Fig. 3.9 b) intensities were quantified using the Image Lab software (Bio-Rad) for the individual Pol III variants. For normalization, band intensities were divided by the total lane intensities for each variant, and the mean values for the replicates were calculated. For illustrative purposes, the derived values were multiplied by a factor 1000. T - termination; RT - Read-through.

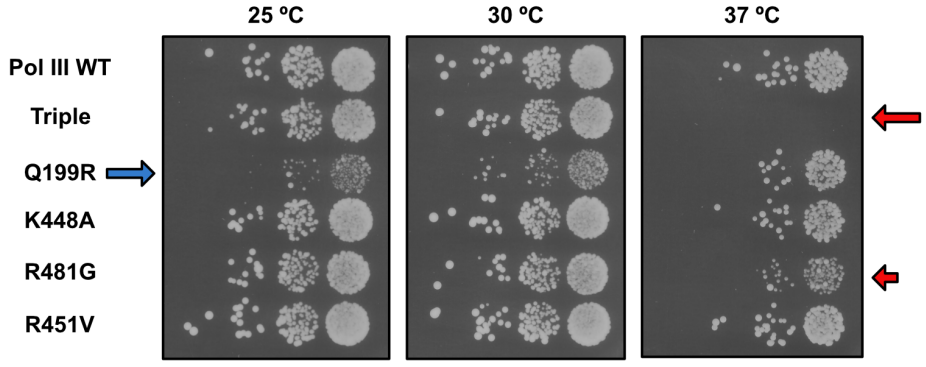

Fig. 3.11. Yeast spotting assay showing the effects of Pol III C128 mutations on yeast growth. Yeast cells were plated as single drops in decreasing concentrations (from right to left) on YPAD plates and incubated at 25 °C, 30 °C and 37 °C for 3 days. Heat-sensitive growth phenotype: red arrows; cold-sensitive growth phenotype: blue arrow. The data shown in this figure was prepared by Dr. Helga Grötsch.

the lab of Dr. Domenico Libri and Dr. Odil Porrua (Université de Paris, CNRS, FR) to analyse the effect on the Pol III C128 mutants via UV cross-linking and analysis of cDNA (CRAC) [186]. CRAC analysis allows detecting Pol III transcription termination RT *in vivo* [149], which can give further insights into how C128 mutations affect Pol III transcription termination. Such analysis could also explain the different growth phenotypes observed in the yeast spotting assays.

3.7. Structure of Pol III Δ gives insights into the function of C53-C37 in Pol III transcription termination

The Pol III PTC structure revealed a key role of subunit C128 in recognizing the poly-dT termination signal. However, it does not explain the role of the C53-C37 heterodimer, which is well-known for being essential for correct Pol III termination [69, 75, 91, 92, 143]. To get insights into the molecular role of C53-C37 in Pol III transcription termination, I purified a Pol III Δ, which lacks the C53-C37 heterodimer and the C11 subunit. The *S. cerevisiae* Pol III Δ strain (kindly provided by Dr. Richard Maraia (NIH, Bethesda, USA)) expresses *S. pombe* C11 as a replacement of deleted *S. cerevisiae* C11 [75]. The C11 substitution permits yeast cell viability but results in loss of C53-C37/C11 during protein purification. The yeast Pol III Δ strain carries an N-terminal 6xHis-FLAG-tagged C128, which allowed me to purify endogenous yeast Pol III Δ via immobilized metal affinity chromatography (IMAC), followed by two successive anion-exchange chromatography steps using a MonoQ 5/50 column (Sigma-Aldrich, Merck). This procedure yielded a protein complex that shows the same protein band pattern as Pol III WT with the exception that the subunits C53-C37/C11 are missing (Fig. 3.12 a). Similarly to the Pol III PTC, I incubated Pol III Δ with scaffold SC1 and solved its structure via cryo-EM (Fig. 3.13 a). The 3D reconstruction had a nominal resolution of 3.9 Å (Fig. 3.13 b) and, locally, reached 3.7 Å in the core (Fig. 3.13 c). Although the overall resolution is lower than of the Pol III PTC, it still allows a side-by-side comparison of Pol III Δ and Pol III PTC bound to the same 7-dT transcription scaffold (Fig. 3.12 b).

As expected, the cryo-EM signal of the C53-C37 heterodimer and subunit C11 are missing in the Pol III Δ map (Fig. 3.12 b). Inspection of the NT-strand binding region reveals that the NT-strand is much poorer resolved in the Pol III Δ structure compared to the Pol III PTC, despite the presence of the 7 dTs on the NT-strand. Although some fragmented cryo-EM density signal for the NT-strand is visible at a lower threshold levels, it still is not good enough to confidently place and refine a model of the NT-strand. Furthermore, a superimposition between the Pol III Δ and PTC structures reveals a relaxation of the Pol III fork loop 1, protrusion and lobe elements (Fig. 3.12 b,c). Thus, the compaction

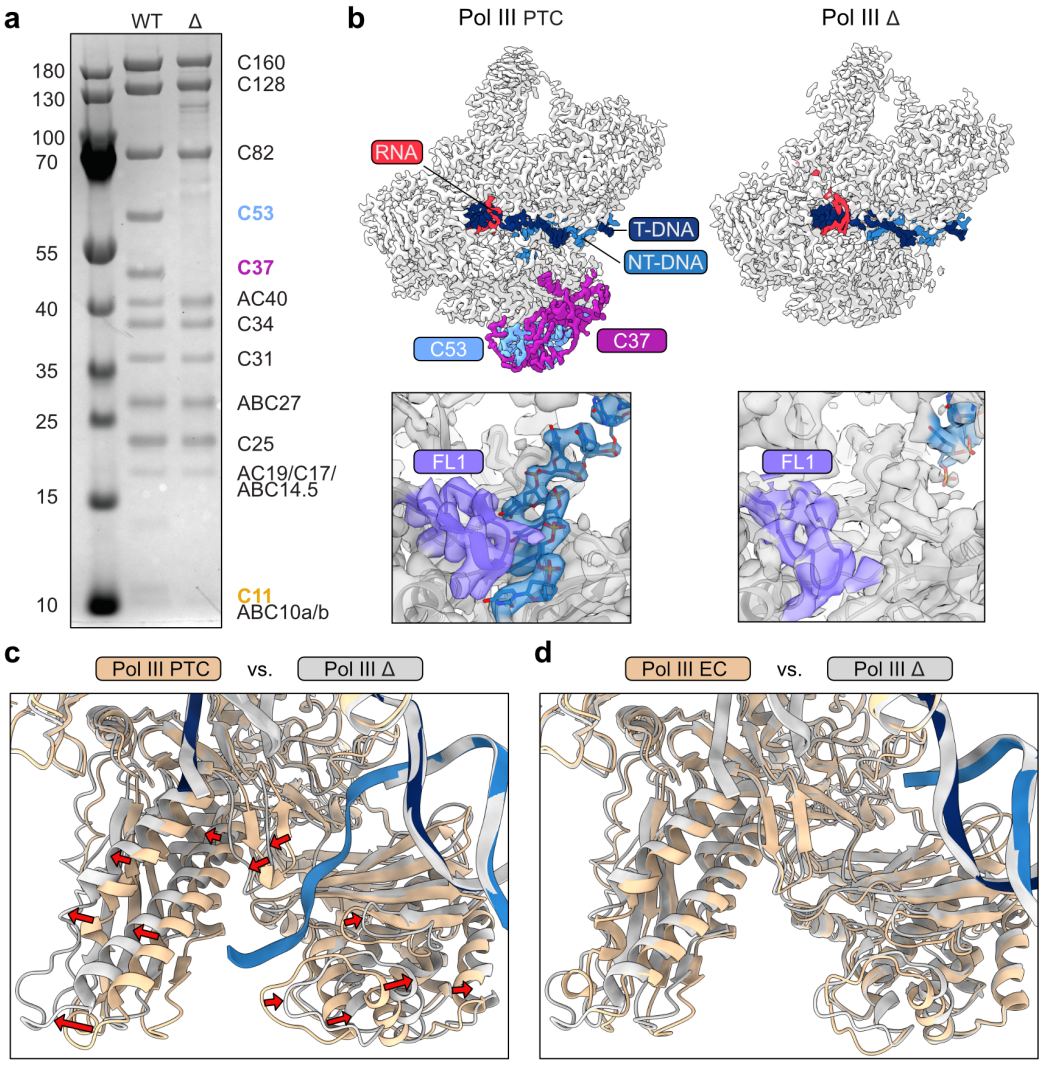

Fig. 3.12. Cryo-EM structure of Pol III Δ bound to the 7-dT transcription scaffold. (a) Coomassie-stained SDS-PAGE of yeast Pol III WT and yeast Pol III Δ. The subunit C53-C37/C11 that are missing in Pol III Δ are highlighted in color. The gel is cropped for clarity. **(b)** Side-by-side comparison of the cryo-EM structures of Pol III Δ bound to scaffold SC1 (carrying 7 dTs) and Pol III PTC. Top: comparison of cryo-EM maps. Bottom: close-up views on the NT-strand (coloured in blue) binding regions. The NT-strand is poorly resolved in the Pol III Δ 3D reconstruction, which, in sharp contrast to the Pol III PTC, prohibits reliable building of NT-strand atomic model. The position of the fork loop 1 (FL1) is different between Pol III PTC and Pol III Δ **(c)** Superimposition of Pol III PTC (coloured) and Pol III Δ (gray). Red arrows illustrate relaxation of C128 in the absence of C53-C37/C11. **(d)** Superimposition of Pol III EC (coloured) and Pol III Δ (gray).

of the Pol III core, which was observed during the transition from Pol III EC to PTC, is, essentially reversed. The Pol III Δ structure, therefore, resembles the Pol III EC structure despite the presence of the 7-dT termination signal (Fig. 3.12 d). Accordingly, the C53-C37 heterodimer, although not contacting the NT-strand directly, indirectly assists Pol III termination by binding C128 in a position that is near to the NT-strand binding region. Based on these findings, I hypothesize that C53-C37 modulates C128 in an allosteric manner. I speculate that the presence of stably-bound C53-C37 reduces the flexibility of C128 and, thereby, reinforces the interaction of C128 and the NT-strand when a poly-dT signal is present. In the absence of C53-C37, C128 might be more flexible and, therefore, less likely to bind the NT-strand in a stable manner. Consequently, Pol III is not slowed down when reaching the termination signal, which abolishes correct transcription termination. Another potential explanation for the essential function of C53-C37 in Pol III termination could be that it tightly interacts with the N-terminal domain (NTD) of subunit C11. Because C11 also plays a critical role in termination, the lack of C53-C37 presumably destabilizes the binding of C11 to the Pol III core, which might also affect transcription termination efficiency.

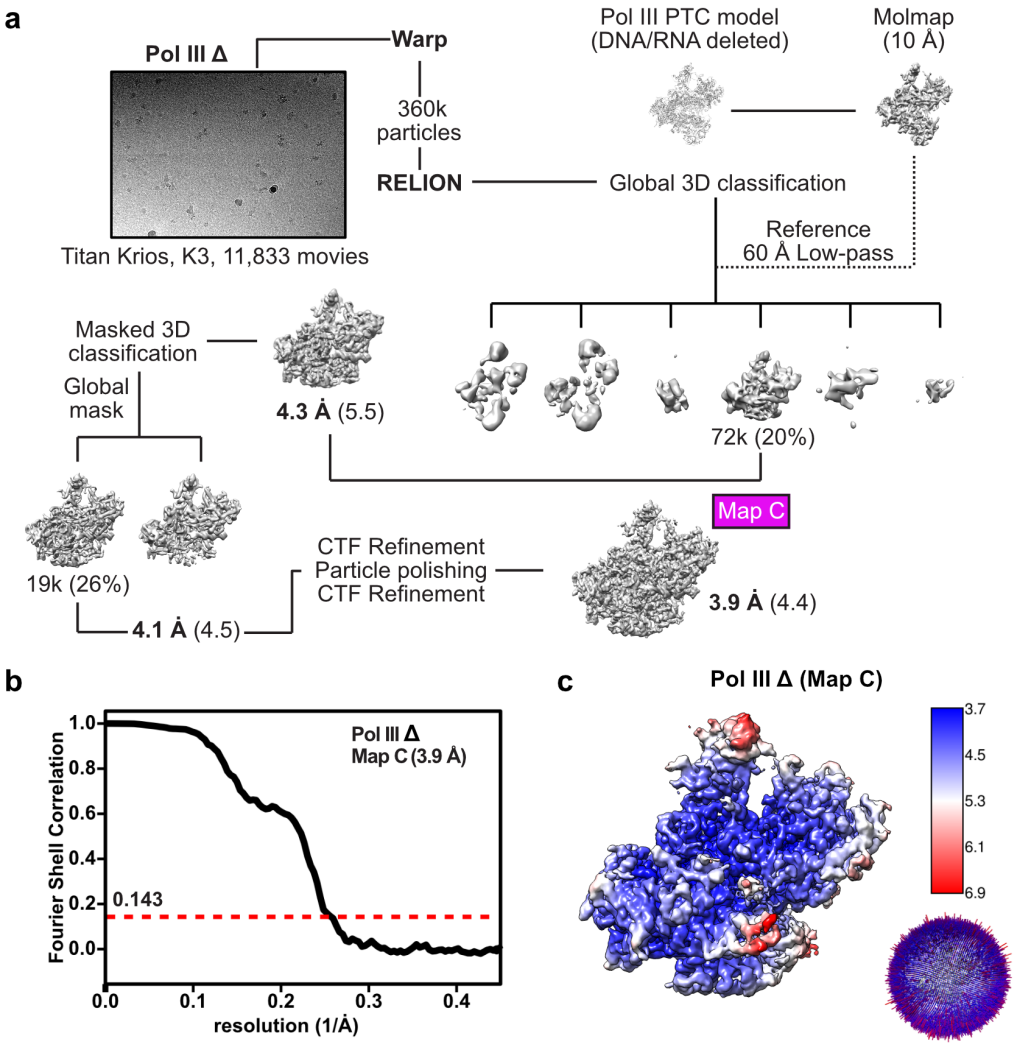

Fig. 3.13. Cryo-EM structure determination of yeast Pol III Δ. **(a)** Data collection and processing strategy of Pol III Δ. Resolution values of sharpened maps are given below and of unsharpened maps in parenthesis. Particle numbers were rounded. **(b)** FSC curve of yeast Pol III Δ (FSC = 0.143). **(c)** Local resolution estimation and angular distribution plot of yeast Pol III Δ.

3.8. Sample preparation of III PTC + NTPs

The primer extension experiments on the scaffold SC1 showed a strong band at the position i+3, which suggested stalling of Pol III after the incorporation of three nucleotides. To better understand why Pol III stalks specifically at the position i+3 and to get further insights into the Pol III transcription termination mechanism, I set out to solve the structure of the yeast Pol III PTC in the presence of NTPs. For initial RNA primer extension experiments, I used the TRB 60 buffer (20 mM Hepes, pH 7.6, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 2 mM DTT); and for cryo-EM sample preparation I used the EB 150 buffer (15 mM Hepes pH 7.5, 150 mM AmSO4, 5 mM MgCl2, 10 mM DTT). I, therefore, first tested if I can also stall Pol III in the presence of NTPs using EB 150. In addition, I tested RNA primer extension using different mixes of NTPs: CTP+ATP+UTP+GTP (C+A+U+G); CTP only (C); CTP+ATP (C+A); CTP+ATP+UTP (C+A+U); and ATP+UTP (A+U). Using the different NTP mixes should reveal the precise position of RNA extension stalling. Fig. 3.14 a shows a schematic of the used transcription scaffold. By using the C+A+U mix, I expected to enrich for the stalled Pol III conformation at position i+3, which should help to obtain a more homogeneous cryo-EM 3D reconstruction. Fig. 3.14 b shows the RNA gel of the RNA primer extension assay. In TRB 60 buffer and in the presence of C+A+U+G, Pol III stalls on position i+3 on SC 1, whereas the corresponding band is absent on SC 2 (compare lane 3 and 10). The usage of C, C+A, and C+A+U further confirms that Pol III stalls on position i+3 (see lane 4, 5, 6), which relies on the correct incorporation of NTPs. When no C is present, as shown by the control experiment (A+U), Pol III does not extent the RNA (lane 7). The assay further reveals that RNA extension is also achieved in EB 150 buffer (lanes 15-18). At higher salt, and in the presence of all NTPs, the stalled band at i+3 is still visible on SC 1 but Pol III can escape the stalled conformation more easily than in TRB 60 as shown by the appearance of longer RNA products (compare lane 3 and 17). Notably, the usage of C+A+U shows a strong enrichment for the position i+3 (lane 18). These results indicate that the C+A+U NTP mix is the preferred choice in obtaining a homogeneous 3D reconstruction, in which Pol III is stalled at position i+3.

Next, I tested which temperature and incubation time would be best suitable for the Pol III PTC + NTP cryo-EM sample. I repeated the RNA primer extension experiment using SC1 as scaffold, EB 150 as reaction buffer and C+A+P as NTP mix. I used a time course from 0 to 90 min and tested three different temperatures: 4 °C, room temperature (RT, ca. 23 °C), and 28 °C (Fig. 3.15 a). At 4 °C, RNA extension activity is low and only little i+3 band products are visible, even after incubation for 90 min. At RT, Pol III is more active and a gradual increase of the i+3 product can be observed over the

time course. Still, after 90 minutes, ca. 30-40% of the RNA primer hasn't been extended yet. At 28 °C, Pol III activity is the highest and a strong enrichment of i+3 can be detected. After 90 minutes, ca. 90% of the RNA primer has been extended to the i+3 position. Thus, incubating the Pol III PTC sample for 90 min at 28 °C in the presence of C+A+U appears to be the best conditions for cryo-EM sample preparation. For the

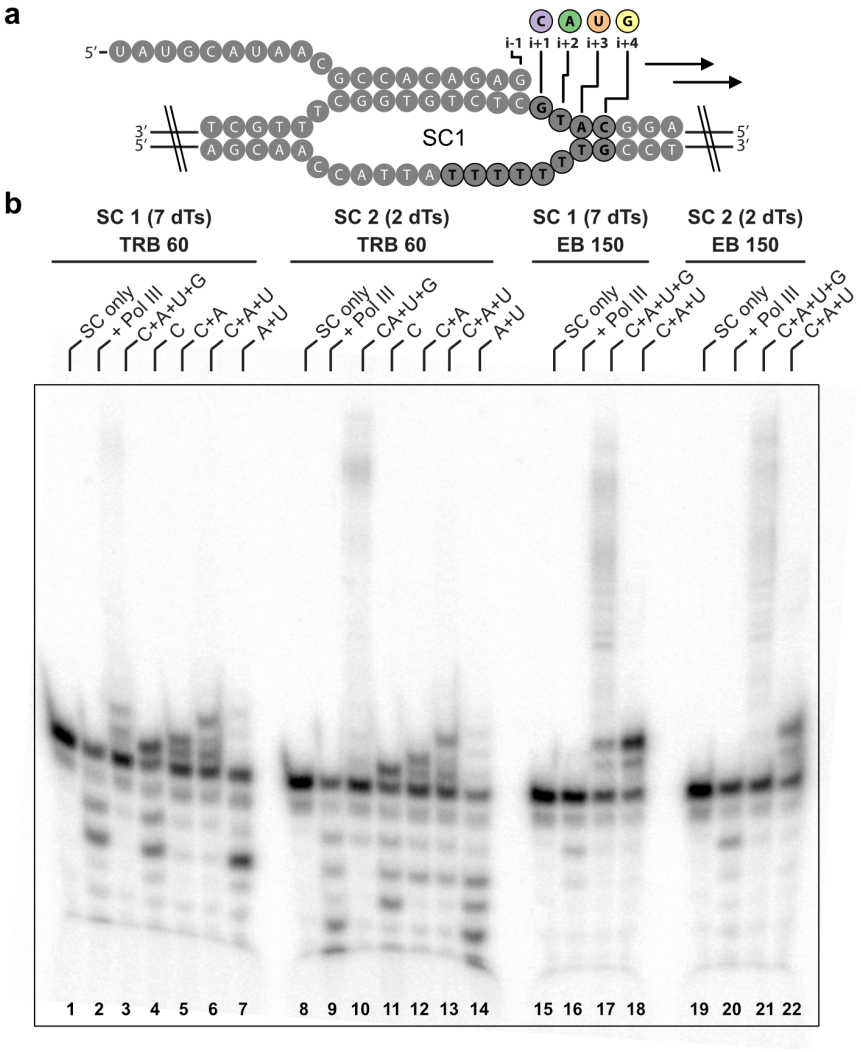

Fig. 3.14. RNA primer extension using different buffers and NTP mixes. (a) Schematic of used transcription scaffold (SC 1) carrying the 7 dT termination signal on the NT-strand. The positions of the NTP incorporation sites (i+N) are labeled and drawn above are the NTPs required to extent the RNA until the respective positions. **(b)** Pol III RNA primer extension experiment for the SC 1 (7 dT) and SC 2 (2 dT) scaffolds. Two reaction buffers and different NTP mixes were used as indicated above lanes. TRB 60 buffer - 20 mM Hepes, pH 7.6, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 2 mM DTT; EB 150 buffer - 15 mM Hepes pH 7.5, 150 mM AmSO4, 5 mM MgCl2, 10 mM DTT; C - CTP; A - ATP; U - UTP; G - GTP. n (Scaffold) = 1.5 pmol; n (Pol III) = 4 pmol; c (NTPs) = 1 mM per NTP. 17% TBE denaturing PAGE. The gel is cropped for clarity.

cryo-EM sample preparation of the Pol III PTC, I supplemented the PTC sample with 4 mM CHAPSO. I, therefore, tested if CHAPSO affects Pol III transcription activity (Fig. 3.15 b). The primer extension experiments on SC 1 in presence or absence of 4 mM CHAPSO shows that CHAPSO does not affect Pol III activity (compare lane 1 and 2). Furthermore, I tested in parallel with a gel shift assay if Pol III is still bound to SC 1 after the incubation with NTPs over 2 h. As shown in Fig. 3.15 c, band shifts can be observed for all tested samples. Moreover, the reaction, in which Pol III was incubated at 28 °C with C+A+U in the presence of CHAPSO shows the strongest signal for Pol III bound nucleic acids (see lane 9). Thus, Pol III remains bound to SC 1 when incubated with C+A+U, even after 2 h incubation time. In summary, by extending the RNA primer extension analysis, I could identify the most suitable conditions for cryo-EM sample preparation of the Pol III PTC + NTPs structure. Hence, to solve the structure of Pol III + NTPs, I incubated Pol III + SC 1 at 28 °C for 90 min with C+A+U in the presence of 4 mM CHAPSO before cryo-EM grid preparation.

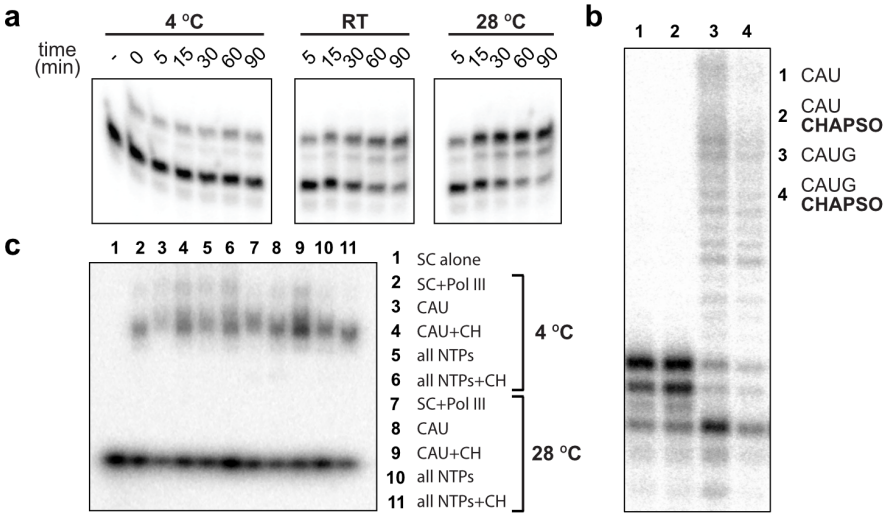

Fig. 3.15. RNA primer extensions to test for best suitable cryo-EM conditions of Pol III PTC + NTPs. (a) Time course of Pol III RNA primer extension on SC 1 in the presence of C+A+U at indicated temperatures. **(b)** Pol III RNA primer extension experiment on SC 1 in the presence or absence of 4 mM CHAPSO. **(c)** Band shift assay of Pol III bound to SC 1 after 2 h incubation with NTP mixes. **(a)-(c)**: Buffer - EB 150; Samples: n (Scaffold) = 1.5 pmol; n (Pol III) = 4 pmol; c (NTPs) = 1 mM per NTP. **(a)-(b)**: 17% TBE denaturing PAGE. **(c)**: 3-8% Tris-Acetate native PAGE in TG buffer. Gels are cropped for clarity.

3.9. Cryo-EM structure of Pol III PTC + NTPs

Collecting a high-resolution data set yielded a 3D reconstruction of the Pol III PTC + NTPs at 2.7 Å nominal resolution (Map D, Fig. 3.16 a and Fig. 3.17 a,b) that extended to 2.5 Å, locally (Fig. 3.17 d). During data processing, it became apparent that the upstream DNA is better resolved than in the Pol III PTC. I, therefore, used masked 3D classification on the DNA and RNA elements and obtained a 3D reconstruction at 3.2 Å, which showed improved cryo-EM density signal for the upstream DNA (Map E, Fig. 3.16 b,c and Fig. 3.17 a,c). The local resolution values of the upstream DNA in Map E ranged between 4.7 and 6.5 Å (Fig. 3.17 e). The cryo-EM signal was of sufficient quality to reveal the minor groove of the upstream DNA (Fig. 3.16 c). This allowed, for the first time, building the model of the complete Pol III-bound transcription bubble (Fig. 3.16 d).

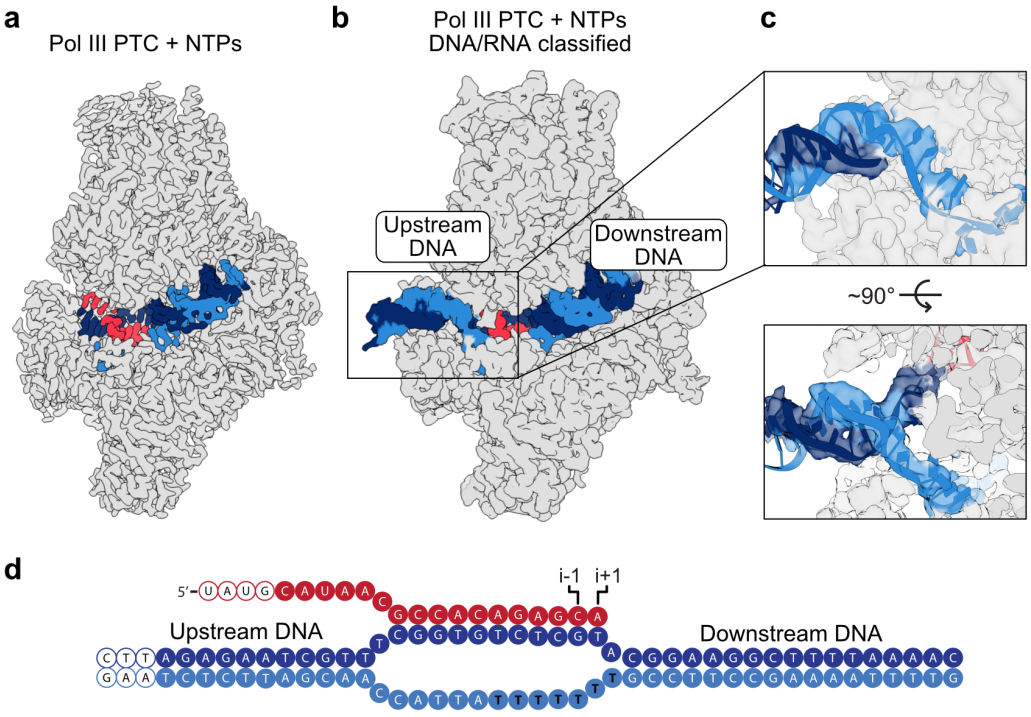

Fig. 3.16. Cryo-EM structure of Pol III PTC + NTPs. (a) Cryo-EM maps of Pol III PTC + NTPs (map D) and **(b)** of Pol III PTC + NTPs, classified on DNA:RNA (map E). **(c)** Close-up views on upstream DNA. **(d)** Schematic of Pol III-bound DNA:RNA scaffold. Filled circles represent built bases.

A close-up inspection of the Pol III active-site in the Pol III PTC + NTPs structure gave insights into why the Pol III adopts a stalled conformation. The cryo-EM densities

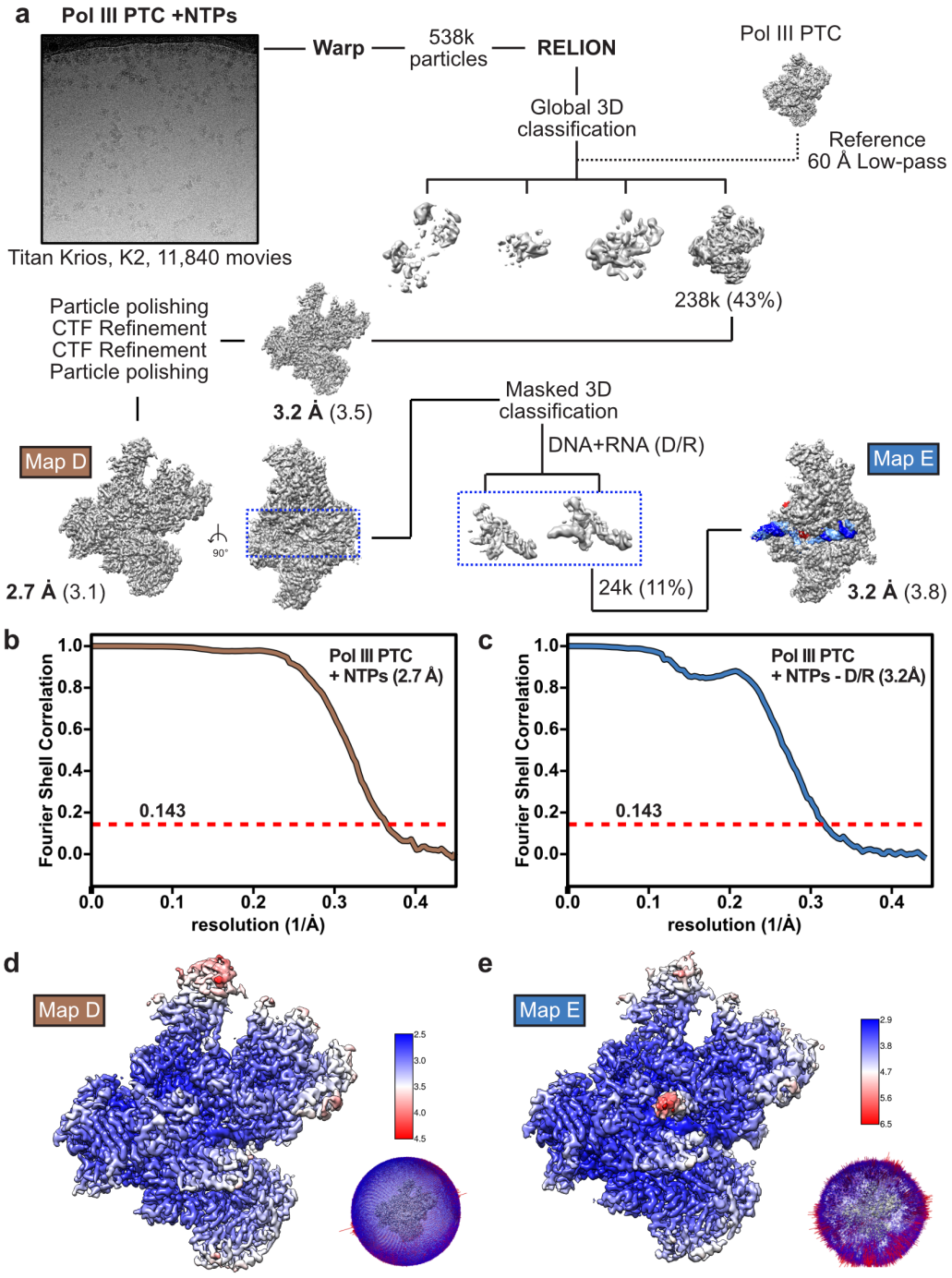

Fig. 3.17. Cryo-EM structure determination of yeast Pol III PTC + NTPs. (a), Data collection and processing strategy of Pol III PTC + NTPs (map D). Masked 3D classification on DNA+RNA (D/R) yielded a map with improved signal for the upstream DNA. Resolution values of sharpened maps are given below and of unsharpened maps in parenthesis. Particle numbers were rounded. **(b)**, **(c)** FSC curves of map D **(b)** and map E **(c)** (FSC = 0.143). **(d)**, **(e)** Local resolution estimations and angular distribution plots of map D **(d)** and map E **(e)**.

of map A (Pol III PTC) and map D (Pol III PTC + NTPs) are of sufficient qualities to identify the registers of the bound nucleic acids relative to the Pol III active sites (Fig. 3.18). In both 3D reconstructions, the DNA:RNA hybrids are well resolved. In the Pol III PTC structure, the guanine (G19) of the 3'-end of the bound RNA primer anneals at position i-1, and the i+1 site is not occupied (Fig. 3.18 a). In the Pol III PTC + NTPs structure, all bases of the DNA:RNA hybrid shifted one position away from i-1 site, the G19 is at position i-2 and the i-1 site is occupied by a cytosine (C20) (Fig. 3.18 b). Thus, the RNA got extended by at least one nucleotide, confirming that the NTP addition experiment to prepare the cryo-EM sample of the Pol III PTC + NTPs structure was successful. In addition, density is visible for a bound or incoming nucleotide (A21) at position i+1. At even lower threshold levels, some extra density appears at the putative position i+2, which could correspond to a UTP that was also present in the NTP mix. Occupation of the i+2 side would require backtracking of the RNA polymerase. However, the cryo-EM density is not of sufficient quality to unambiguously build a nucleotide inside, which is why the final model of the Pol III + NTP structure does not contain an UTP at the 3'-end. The weak incorporation or occupation of the UTP at the 3'-end of the RNA primer is surprising because the primer extension experiments suggested a clear enrichment of the RNA at the position i+3, which would require the incorporation of the UTP. A potential explanation is that the UTP got attached to the RNA but the 3'-end of the RNA at the position i+2 may only loosely associate with the polymerase. Hence, the i+2 side might be too flexible to be well resolved in the 3D reconstruction.

In both 3D reconstructions, the NT-strand is well resolved and all 7 dTs of the termination signal are clearly visible. Interestingly, the register of the NT-strand in the Pol III PTC + NTPs structure didn't change compared to the Pol III PTC structure. Thus, addition of NTPs induces extension of the RNA primer and a register shift of the DNA:RNA hybrid but leaves the NT-strand unchanged. This observation reinforces the model that the tight interaction between C128 and the NT-strand in the Pol III PTC is sufficient to pause Pol III, which is a prerequisite for subsequent termination. In addition, the register offsets between the NT-strand and the DNA:RNA hybrid may augment stalling of Pol III and traps Pol III in an unproductive state. Adopting such an unproductive conformation could explain how termination is finally triggered. I hypothesize that the TFIIS-like subunit C11 senses this stalled conformation and, in attempts to rescue Pol III from this conformation, induces transcription termination. I will elaborate on this hypothesis in the next section.

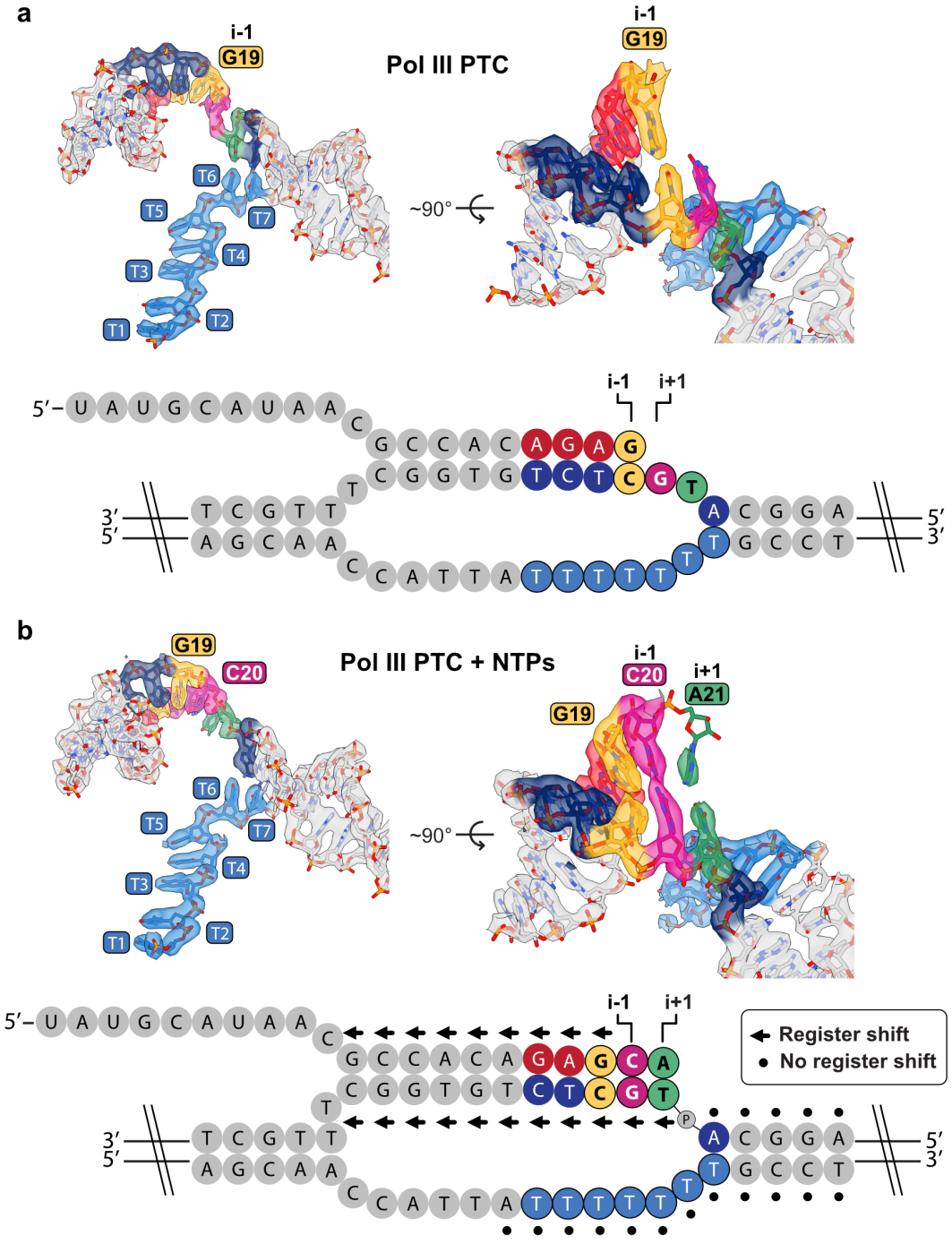

Fig. 3.18. Addition of NTPs induces register shift of the T-strand but not of the NT-strand. Close-up views on the nucleic acids in the active site of the Pol III PTC **(a)** and Pol III PTC + NTPs **(b)**. Schematics of the bound DNA:RNA scaffold SC1 are shown below. Coloring highlights different nucleic acid elements or relative positions of the nucleotide addition sites (i). Addition of NTPs induces a register shift (arrows) of the T-DNA:RNA hybrid but not of the NT-strand (dots). Bases without shift labels do not exhibit sufficient cryo-EM density to unambiguously assign their register.

3.10. Model of Pol III transcription termination

By combining the here presented structural and biochemical work with published data, I propose a mechanistic model of Pol III transcription termination and of the Pol III transcription cycle (Fig. 3.19).

Step 1 - Transcription initiation and elongation complex formation: Pol III is recruited to the promoter DNA by general transcription factors (TFs) such as TFIIIB and TFIIIC. Recruitment also involves the Pol III-specific C82-C34-C31 heterotrimer and the C53-C37 heterodimer. The dsDNA is melted into the T-DNA strand and the NT-DNA strand, and an initial RNA primer is synthesized that anneals to the T-DNA strand. The complete transcription bubble forms, Pol III enters the elongation stage and synthesized the RNA. In the Pol III elongation complex, the NT-strand only loosely associated with the Pol III core and is, therefore, to flexible to be visualized in the cryo-EM 3D reconstruction.

Step 2 - NT-strand recognition and PTC formation: Pol III reaches its termination site that harbours at least 5 dT bases on the NT-strand, and the Pol III PTC forms. Subunit C128 recognizes the dTs by forming multiple H-bonds with the dT bases, for which it uses a set of highly conserved residues that are mostly unique to Pol III. Thereby, the C128 lobe, protrusion and fork elements locally contract around the NT-strand and form a positively charged groove. The NT-strand becomes ordered and is threaded through the positively charged groove, and its path is defined by several key residues. In addition, the fork loop 1 element forms a hydrophobic cavity that accommodates the methyl groups of the dT bases. Thereby, Pol III can potentially distinguish between dTs and dCs. Combined, the H-bond network and the binding of the dT methyl groups induce pausing of Pol III, which is a prerequisite for transcription termination [144]. C128-mediated recognition of the NT-strand is assisted by the C53-C37 heterodimer. Although C53-C37 does not bind the NT-strand directly, it stably associates with C128 at a nearby region. This interaction is required for C128-mediated NT-strand recognition because C128 fails to contract around the termination signal and does not stabilize the NT-strand in the absence of C53-C37 as shown by the cryo-EM structure of Pol III Δ. I hypothesize that C53-C37 binding increases the rigidity of C128 and, thereby, forces C128 to interact with dT bases once Pol III reaches the termination signal. Conversely, C128 is presumably more flexible in the absence of C53-C37 and might not engage that tightly with the dTs in the absence of C53-C37 because C128 has more "space to breathe". Consequently, Pol III would fail to pause and subsequently terminate on the poly-dT termination signal when C53-C37 is missing. Thus, C53-C37, albeit being stably integrated into the Pol III structure, modulates Pol III activity in an allosteric manner

once the termination signal is reached. This mode of action of C53-C37 explains the well known role of the C53-C37 heterodimer in Pol III transcription termination [69, 75, 91, 92, 143], which was, however, only poorly understood on a molecular level.

Step 3 - NTP incorporation, nucleic acid register offset and further stalling: Although Pol III is paused via the tight interaction between C128 and the NT-strand, NTPs can still be incorporated into the nascent RNA strand. The strong association between Pol III and the NT-strand, however, hinders translocation along the DNA. Consequently, the NT-strand register remains unchanged whereas the register of the T-strand and the RNA shifts upon addition of NTPs. The register offset potentially further provokes stalling of Pol III and traps it in an arrested state.

Step 4 - Insertion of C11 into Pol III active site: The arrested state could, potentially, be sensed by the TFIIS-like subunit C11, which also functions in Pol III transcription termination [69, 85]. Similarly to TFIIS, C11 harbours an acidic hairpin on its C-terminal domain (CTD), via which it hydrolytically cleaves RNA on its 3'-site. In the Pol II system, TFIIS can reactivate backtracked Pol II that stalled upon NTP misincorporation by cleaving off the backtracked RNA, which protrudes out of the polymerase funnel domain [89]. I speculate that, similarly to TFIIS, C11 senses the arrested conformation of Pol III. The register offset and stalling of Pol III may induce backtracking of Pol III and protrusion of the backtracked RNA out of the Pol III funnel. In attempts to rescue the arrested state, the CTD of C11 may insert into the Pol III funnel to reach the active site and to cleave of the backtracked RNA.

Step 5 - C11-mediated RNA cleavage and clamp opening: The insertion of C11 potentially induces partial opening of the Pol III clamp domain, which contacts the downstream DNA. Evidence for the opening of the clamp comes from my work on the human Pol III structure, where 3D variability analysis on the human Pol III EC revealed opening of the clamp upon insertion of RPC10 (the human ortholog to yeast C11) into the Pol III funnel (see Fig. 2.9). C11 will cleave the 3'-end of the RNA, which has been reported to occur concomitantly to transcription termination [144] but is not a strict prerequisite for termination to take place per se [69].

Steps 6 and 7 - Destabilization of the DNA:RNA hybrid and subsequent termination: The Pol III termination signal consists of at least 5 dT bases on the NT-strand in *S. cerevisiae*. Consequently, the corresponding complementary T-strand contains a poly-adenine (dA) stretch, and a poly-adenine:uracile (dA:rU) DNA:RNA hybrid forms in the active site during RNA synthesis. Importantly, dA:rU hybrids are remarkably unstable compared to other DNA:RNA hybrids [147]. Thus, the poly-dA:rU hybrid that forms once Pol III reaches the termination site is more likely to collapse.

During transcription elongation, the DNA:RNA hybrids is an important stabilizer of the RNA polymerase elongation complex [146]. The collapse of the unstable poly-dA:rU hybrid could, thus, likely serve as the final signal that triggers Pol III transcription termination. The collapse of the hybrid and destabilization of the Pol III-nucleic acid complex could be further provoked by the insertion of C11 and the concomitant clamp opening. In the Pol III EC, the DNA:RNA hybrid is more loosely bound to the active site compared to Pol I and Pol II [16, 148]. In contrast, the Pol III clamp tightly associates with the downstream DNA, which may compensate for the loose binding of the DNA:RNA hybrid in the active site. The opening of the clamp domain upon C11 insertion into the active site could interfere with the tight grip of the downstream DNA and, therefore, with its compensatory stabilizing role during transcription elongation. Combined, these effects, presumably trigger the final step of transcription termination; namely, the collapse of the Pol III transcription complex and the release of the DNA and the RNA transcript.

Step 8 - Facilitated recycling of Pol III: Once the RNA transcript got released, Pol III can be recycled to the promoter DNA to initiate a new round of RNA transcription. Pol III, once terminated, can undergo facilitated recycling that is much faster than the initial transcription initiation [135]. C11 has been shown to be critical for the facilitated recycling of Pol III [69]. Structural data of the Pol III pre-initiation complex (PIC) suggested that opening of the Pol III clamp is required to allow binding to the closed promoter DNA [77, 80]. Hence, I speculate that C11-mediated opening of the clamp domain not only triggers transcription termination but also assists facilitated recycling of Pol III. In this regard, terminated apo Pol III may retain C11 inserted into the funnel domain. Consequently, such Pol III variant could feature an open clamp domain that is more capable of binding and melting the closed promoter DNA.

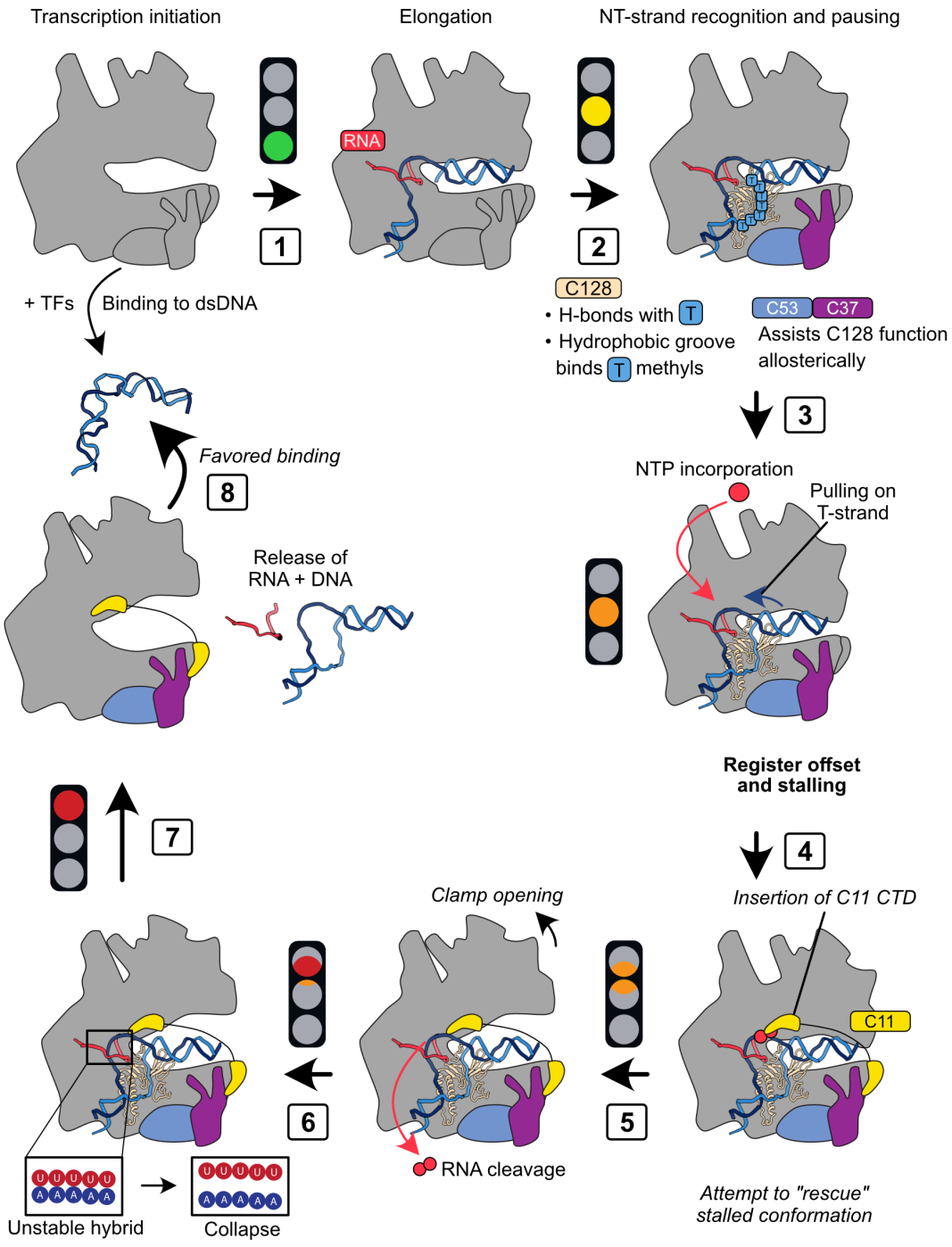

Fig. 3.19. Mechanistic model of Pol III transcription termination. Model of Pol III transcription termination combining here presented structural and biochemical data with evidence from the literature. For detailed description, see main text. Literature-based hypotheses are italicized.

3.11. Transcription termination of human Pol III

Pol III transcription termination in yeast requires at least 5 dT bases on the NT-strand. Interestingly, most human tRNA genes (69%) contain only a 4 dT termination site [184], which raises the question of how human Pol III terminates transcription efficiently. It has been suggested that Pol III might also rely on auxiliary factors to terminate transcription, which, particularly could be the case for human Pol III given that most human tRNA genes contain less than 5 dTs [187]. Although transcription termination of human Pol III on 4 dTs has been reported [188], it remained unclear if human Pol III can terminate transcription on its own because the cited study used nuclear extracts of HeLa cells, which could also contain the required auxiliary factors. To test, if human Pol III is able to autonomously terminate transcription on a 4 dT termination site, I took advantage of the fact that we can work with purified and highly homogeneous human Pol III sample. I designed a tailed template transcription assay using the sequence of a human tRNA gene tRNA-Arg-CCG-1-3 that has been reported to be upregulated in breast cancer cells [189], and which contains the following termination sequence: TTTTCTT (Fig. 3.20 a). The experiment was performed by Dr. Florence Baudin who used endogenous human Pol III and also tested yeast Pol III, which enabled a site-by-site comparison between the two Pol III variants (Fig. 3.20 b). Testing the canonical 4 dT termination sequence confirmed that human Pol III terminates on 4 dTs although some read-through products were detectable (lane 2). In sharp contrast, yeast Pol III completely reads through the canonical human tRNA termination site (compare lane 2 and 4), thereby revealing a clear difference in the activities of the two polymerases. When the dC at position 5 was replaced with another dT (resulting in a 7 dT termination signal), both human and yeast Pol III terminated correctly (compare lane 6 and 8). However, read-through products could be detected for yeast Pol III (ca. equal amounts compared to termination products) whereas no read-through transcript are visible for human Pol III. This observation shows that human Pol III is more sensitive to intrinsic termination than yeast Pol III. The third construct that was tested by Dr. Baudin contained the following sequence: TTTTGGT. The dC5 and dT6 were replaced with dGs to analyse if human Pol III relies on a stretch of 5 or more pyrimidines (as present in the original termination sequence) or if only 4 dTs are sufficient to trigger transcription termination. Testing this construct showed that human Pol III terminates as efficient on the substituted termination sequence as on the canonical terminator (compare lane 2 and 10). In summary, this experiment confirmed that human Pol III is able to terminate autonomously on 4 dTs whereas yeast Pol III requires longer poly-dT termination sequences.

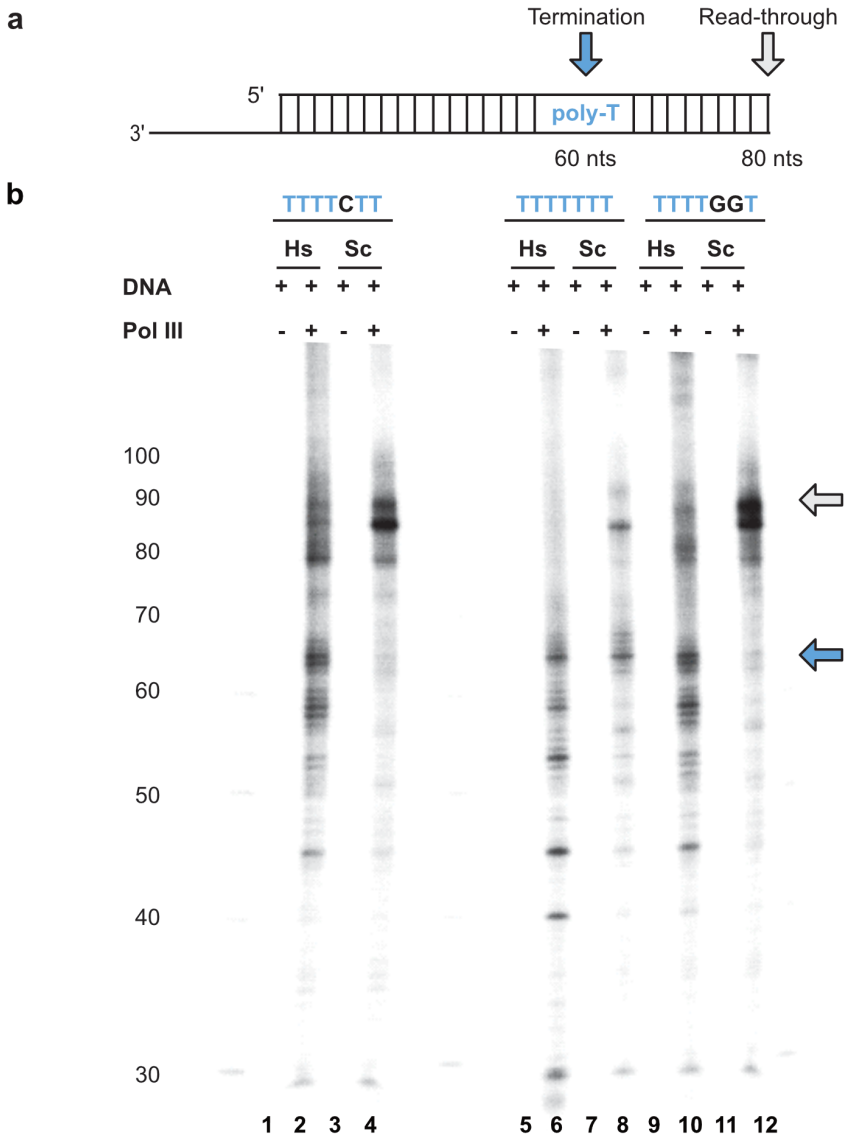

Fig. 3.20. Transcription termination analysis of human Pol III. (a), Schematic of the tailed-template transcription construct based on the human tRNA-Arg-CCG-1-3 gene. The termination sequence is highlighted with a blue arrow. The grey arrow marks potential read-through products that would be released once Pol III reaches the end of the dsDNA. **(b)**, Denaturing RNA gel of the tailed-template transcription assay. Three constructs with varying termination sequences were tested as shown above the lanes. Hs - *H. sapiens*; Sc - *S. cerevisiae*. n (Pol III) per lane: 2 pmol; n (DNA) per lane: 1 pmol. 15% TBE-Urea PAGE. The experiment was performed by Dr. Florence Baudin.

These clear differences raise the question why human Pol III is capable to terminate on shorter dT tracts. Most residues that contact the DNA:RNA hybrid in the Pol III active site are either identical or conserved between human and yeast Pol III (Fig. 2.7 a). It is, therefore, unlikely that the increased sensitivity of human Pol III for poly-dTs tracts is a consequence of a more loosely bound DNA:RNA hybrid. Instead, I speculate that the differences in termination activities could be explained by different activities of C11 (yeast) and RPC10 (human). In the human Pol III EC structure, ca. 25% of the particles adopted the RPC10 "inside funnel" conformation (Fig. 2.4 and Fig. 2.8). A close-up inspection of the active site showed that human Pol III was still in an actively transcribing conformation despite the insertion of RPC10 into the funnel domain. Thus, human RPC10 is able to actively monitor the active site during RNA synthesis. In case of yeast Pol III, no insertion of the C11 CTD has been observed yet, suggesting that yeast C11 only inserts into the funnel once Pol III stalls or backtracks, which might also occur in a more transient manner. The higher tendency of human RPC10 to insert into the funnel and its capability to monitor the active site could serve as an intrinsic destabilizer that renders human Pol III more sensitive poly-dT termination signals. However, termination of human Pol III has to be studied in more detail (which could involve the structural analysis of the human Pol III PTC) to better understand the observed differences between human and yeast transcription termination.

3.12. Transcription termination - conclusions and outlook

In this section, I have presented the structural basis that gives rise to Pol III transcription termination. The cryo-EM structure of the yeast Pol III PTC revealed how Pol III recognizes the poly-dT termination signal on the NT-strand. Subunit C128 forms multiple H-bonds with the NT-strand via a set of residues, which are mostly unique to Pol III. In addition, the fork loop 1 element forms a hydrophobic cavity that accommodates the methyl groups of the dT bases, via which dT and dC bases can be distinguished. These interactions induce pausing of Pol III that is necessary for the subsequent termination event, which I could confirm by structure-function studies.

The structure of Pol III Δ shed light on the function of the C53-C37 heterodimer in transcription termination. Unlike previously proposed by our group [16], subunit C37 does not directly contact the NT-strand in the PTC conformation. The Pol III EC structure revealed that subunit C37 binds the C128 lobe. Two helical elements (184-203 and 220-240) bind C128 close to the DNA binding cleft and are separated by a flexible loop, which was not resolved in the Pol III EC structure [16]. Based on this observation, it was suggested that the flexible loop recognizes the termination signal and becomes

ordered in the context of the Pol III PTC. This, however, is not the case in the structures that I have presented here. In the Pol III PTC, the C37 loop remains disordered and only C128 binds the NT-strand. Instead, the C53-C37 heterodimer seems to function in Pol III transcription termination in a more indirect manner as shown by the Pol III Δ structure. In the absence of C53-C37, the poly-dT termination signal does not become ordered and the compaction of the C128 lobe, protrusion and fork does not occur. Thus, binding of C53-C37 to C128 enables this subunit to interact with the NT-strand. I speculate that the binding reduces the intrinsic flexibility of C128. Once Pol III reaches the termination signal, C128 is, essentially, forced to tightly engage with the NT-strand. In the absence of C53-C37, C128 may have more molecular freedom and might not form the tight H-bond network that is required for Pol III pausing. This model might, however, be difficult to experimentally validate further. One possibility could be to conduct a molecular dynamics (MD) analysis using the available Pol III structures to, computationally, test if C53-C37 indeed reduces the intrinsic flexibility of C128.

To get insights into the sequence conservation of the key residues that function in NT-strand recognition, I aligned the sequences of C128 orthologs that covered a broad range of the eukaryotic tree [190]. 9 out 14 residues show a high-conservation score (0.8-1.0) across the 5 eukaryotic supergroups that were conserved in the MSA. This suggests that these residues were already present in the Pol III of the last eukaryotic common ancestor (LECA). One of these residues, Q199, is identical in all 15 analysed species and was part of the structure-function studies that I performed. We mutated residue Q199 to an arginine, which can be found in both Pol I and Pol II at the same position in the MSA of A135 (Pol I), RPB2 (Pol II) and C128 (Pol III). This substitution severely reduced the termination efficiency *in vitro* and showed a clear growth phenotype *in vivo*. The high conservation of Q199 and its key role in Pol III termination suggests that LECA's version of Pol III was already capable of terminating on poly-dT stretches. To validate this theory further, a more comprehensive phylogenomic analysis would, however, need to be conducted, which goes beyond the scope of this thesis.

Transcription termination is an essential process across the entire tree of life. However, termination is also the least understood step during the transcription cycle. The main reason why our knowledge about transcription termination is more fuzzy than about transcription initiation and elongation lies in the transient nature of this process. During transcription initiation and elongation, different factors need to assemble into a stable functional unit. Hence, the characterization of these functional states requires the reconstitution of a stable complex. Although chasing such stable complexes can be very challenging and might also fail, the success of the experiment means that the experimental outcome permits the structural analysis of the state of interest. The opposite is true for

transcription termination, for which the transcription machinery must be destabilized. A successful experiment, therefore, means to reconstitute a state that is, per definition, highly unstable and more difficult to capture. Eukaryotic Pol I and Pol II, as well as bacterial and archaeal RNA polymerases require additional factors to terminate [187]. Reconstituting a termination-like state, therefore, requires the addition of factors that destabilize the system. In contrast, for Pol III to terminate *in vitro*, no external factors are needed besides the termination signal on the DNA and the polymerase itself. This unique feature enabled me to propose a structure-based model of the Pol III termination process.

A common feature of transcription termination is the pausing of the polymerases, which might be a requirement for subsequent termination [134]. However, pausing of eukaryotic RNA polymerases at termination sites is not well understood. In this work, I gave an in-depth structure-based picture of how Pol III pausing is achieved on its termination sites. These insights also enable the comparison with the other eukaryotic RNA polymerases. Many of the residues that bind the NT-strand in Pol III and are required for Pol III pausing, lie in the fork loop 1 and fork loop 2 elements, which can be also found in Pol I and Pol II. Interestingly, the fork loop 1 adopts a different folds and conformations when comparing the three polymerases. In Pol I, fork loop 1 is much wider and would clash with the NT-strand when the Pol I and Pol III structure are superimposed. This indicates that the NT-strand needs to run along a different path in Pol I and that fork loop 1 plays a different role in the context of Pol I transcription. In agreement to this, 13 out of 14 residues that contact the NT-strand in Pol III, are not conserved between C128 (Pol III) and its paralog A135 (Pol I). The comparison between Pol III and Pol II shows that some residues in C128 (K230, K393, R446, R451) that bind the NT strand can also be found in RPB2. Interestingly, the fork loop 1 of Pol II, which contains two of these conserved residues (K470, R476), points into a different direction and does not bind the NT-strand. Instead, it contacts the DNA:RNA hybrid, which may have an additional stabilizing function in the Pol II EC. Hence, the mechanism of transcription pausing is, most likely, very different between Pol III and the other two RNAPs. My data suggests that two structural elements that are present in all three RNAPs, the fork loop 1 and fork loop 2, were fine-tuned in the Pol III structure to specifically bind the NT-strand and pause Pol III so that it can terminate without any co-factors.

Interestingly though, the RNA and DNA helicase Sen1 - which is a well-known Pol II termination factors - has recently been shown to be also involved in transcription termination of *S. pombe* Pol III [150]. Together with the RNA-binding proteins Nrd1 and Nab3, Sen1 participates in Pol II transcription termination of non-coding RNAs [191]. In the Pol II system, Sen1 translocates along the nascent RNA chain and triggers

termination by dissociating a paused Pol II complex in an ATP-dependent manner [192]. Using chromatin immunoprecipitation (ChIP) combined with next-generation sequencing (ChIP-Seq) in an Sen1-deleted cell line, Rivosecchi et al. could show that deletion of Sen1 increases transcription read-through [150]. Independently from this work, the lab of Dr. Domenico Libri and Dr. Odil Porrua (Université de Paris, CNRS, France) obtained similar results in *S. cerevisiae* using UV crosslinking and analysis of cDNA (CRAC) (personal communication with the Libri lab). In collaboration with our lab, the Libri lab could show that Sen1 mainly acts on stalled Pol III molecules that failed to terminate on weak primary termination sequences containing only short dT motifs (Xie et al, unpublished). Whereas this process appears to function in an Nrd1 and Nab3-independent manner, it seems likely that Sen1, similarly to its role in Pol II termination, dissociates Pol III transcription complexes via translocation along the RNA and ATP hydrolysis. It is noteworthy that the exact mechanism of how Sen1 dissociates Pol II or Pol III transcription complexes is not known. Future work as part of a planned collaboration between the Libri lab and ours will, hopefully, shed light on the precise mechanism of how Sen1 assists in Pol III transcription termination. Despite the findings that an additional factor, Sen1, can participate in Pol III transcription termination, Pol III still seems to be able to terminate transcription on its own. Furthermore, the model that Sen1 functions as a fail-safe termination factor of Pol III is complementary to my findings. Sen1 functioning would still require Pol III stalling on secondary downstream poly-dT stretches, the mechanism of which has been thoroughly described in this work.

Part III.
Material and methods

4. Material and methods

4.1. Molecular Biology

4.1.1. Polymerase Chain Reaction

For molecular cloning, genes or plasmid backbones were amplified by Polymerase Chain Reaction (PCR) in 50 µL reactions (Table 3).

Table 3. PCR with Phusion DNA polymerase. Left: PCR reaction composition, Right: PCR program. For annealing the same melting temperature (T_M) as for the primer was used. For elongation, 15-30 s per kbp were used.

Component	Volume [µL]	Temperature [°C]	Time	Cycle
5x Phusion HF buffer	10			
12 µM fw primer	1.25	98	30 s	1x
12 µM rvprimer	1.25	98	10s	
10 mM dNTPs	1	T_M (Primer)	30s	35x
Template DNA	1	72	Varies	
Phusion	0.5	72	10 min	1x
Water	35			

4.1.2. Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA molecules according to their size. Typically, 1% agarose gels were cast in 1x TBE and 5 µL DNA Gel Stain SYBR Safe (Life Technologies) were added to stain DNA. Agarose gels were run in 1x TBE buffer for 30 minutes at 110 V. DNA fragments were visualised with a Molecular Imager Gel Doc XR+ Imaging System (Bio-Rad) and a XcitaBlue Screen by recording fluorescent signal at a wavelength of 470 nm.

4.1.3. Chemical transformation of *Escherichia coli* cells

For transformation, 70 µL chemically competent XL1-Blue *Escherichia coli* cells were thawed on ice and 1 µL of plasmid DNA or 10 µL of Gibson cloning reaction were added. The mixture was incubated for 22 minutes on ice and heat-shocked for 45 seconds at 42 °C. The cells were incubated on ice for 2 minutes and 300 µL LB medium were added. Cells were shaken at 37°C at 200 revolutions per minute (RPM) for 1 hour. 200 µL were streaked out on LB plates containing a resistance marker. If ampicillin was used, cells were directly plated without recovery.

4.1.4. DNA isolation from *E. coli* cells

For DNA isolation, 6 mL LB plus antibiotics were inoculated with single colonies and incubated over night at 37°C and 200 RPM. The next day, bacteria were harvested by centrifugation for 7 minutes at 4000 RPM and the DNA was isolated with a plasmid isolation miniprep kit (Qiagen).

4.1.5. Gibson assembly

Gibson assembly [176] was used for molecular cloning. Genes and target vectors were amplified via PCR, and primers were designed to add a minimum of 20 base pairs complementary overhang to its target DNA. 1 µL of DpnI (NEB) was directly added to the finished PCR reaction to digest remaining plasmid DNA for 60 minutes at 37 °C . The PCR products were purified by agarose gel electrophoresis, DNA bands were excised, and the DNA was isolated via a DNA gel extraction kit (Qiagen). 100 ng of the linearized target vector were mixed with gene insert DNA in a 5-fold molar excess. Up to 5 µL of the DNA mixture was added to a 15 µL isothermal reaction aliquot (5% PEG-8000, 100 mM Tris-HCl pH 7.5, 10 mM MgCl2, 10 mM DTT, 0.2 mM dNTPs, 1 mM NAD, T5 Exonuclease (1:1000), Phusion polymerase (1:60), Taq ligase (1:7.5)), the reaction was filled up to 20 µL final volume with water, incubated for 30 minutes at 50 °C, and transformed into XL1-Blue *E. coli* cells.

4.1.6. Cloning of recombinant human Pol III expression constructs

Human Pol III-coding genes were ordered as cDNA clones (see Table A1) and cloned into the pcDNA3 vector in between the cytomegalovirus (CMV) promoter and the Bovine growth hormone polyadenylation (BGH-PolyA) signal with Gibson assembly. Multiple-gene expression constructs of human Pol III were cloned via biGBac cloning [177] with a set of self-made cloning plasmids that were generated according to [177] using pACEBac1 as a starting vector. The biGBac cloning approach was adapted for mammalian protein expression by using a set of modified Cas primers that annealed to the CMV promoter (forward primers) and the BGH-PolyA signal (reverse primers). First, human Pol III genes were amplified by PCR using the modified Cas primers and gel purified to create gene expression cassettes (GECs) that harbour complementary ends to other GECs. In a second step, up to five GECs were cloned into a pBig1 plasmid to create polygene cassettes (PGCs). The desired GECs were mixed in a 5-fold molar excess with 100 µg of SwaI (NEB) linearized pBig1 vectors, the GECs were joined via Gibson assembly, following transformation and DNA isolation. Third, PGCs were joined together into a

pBig2 vector by mixing the PGC-containing pBig1 plasmids in a 5-fold molar excess with 33 ng of PmeI (NEB) linearized pBig2. Water was added to volume of 12.5 µL and 4 µL of 5x isothermal reaction buffer without enzymes were added. The PGCs were released by adding 1 µL of PmeI, and the reactions were incubated for 90 to 120 minutes at 37 °C . The mixture was placed on ice and 2.5 µL of a Gibson enzyme mix (8 µL *Taq* DNA Ligase, 1 µL Phusion DNA Polymerase, 1 µL T5 Exonuclease (prediluted 1:30) were added. Gibson assembly was performed as described above. Gentamycin and chloramphenicol were used as selective markers for cloning into pBig1 and pBig2 plasmids, respectively. After each Gibson assembly step, the generated isolated plasmids were test-digested with SwaI to screen for successful assembly, followed by Sanger sequencing (Eurofins Genomics) for pBig1 plasmids and full-plasmid next-generation sequencing (MGH CCIB DNA Core) for selected pBig2 plasmids to confirm successful cloning.

4.1.7. Mutagenesis of yeast Pol III

Subunit C128 of yeast Pol III was genomically edited via CRISPR-Cas9 by Dr. Helga Grötsch to generate termination-deficient Pol III variants in the following way: Guide RNAs (sgRNA) were designed using the CRISPR tool implemented into Benchling Biology Software and ordered as two short DNA oligonucleotides (forward and reverse directions). For sgRNA assembly, the two oligos were phosphorylated on the 5'-ends with T4 PNK (NEB), annealed in a thermocycler and cloned into the pWS082 sgRNA entry vector, provided by the Steinmetz lab (EMBL Heidelberg) and a gift from Tom Ellis (Addgene plasmid # 90516). Donor DNAs were designed to harbour 50 bp homology on either side of the estimated dsDNA breaks and ordered as two partially (20 bps) complementary DNA oligonucleotides. The donor DNA oligonucleotides were mixed, amplified via PCR and purified with ethanol precipitation. Prior to transformation into yeast cells, the Cas9-sgRNA gap repair vector pWS174, provided by the Steinmetz lab (EMBL Heidelberg) and a gift from Tom Ellis (Addgene plasmid # 90961) was digested with BsmBI (NEB), and the cloned sgRNA entry vectors were linearized with EcoRV (NEB) and gel purified. Next, 100 ng of linearized Cas9-sgRNA gap repair vector, 200 ng of linearized sgRNA entry vector and 5 µg of annealed donor DNAs were transformed with sorbitol into a *S. cerevisiae* strain harbouring c-terminally TAP-tagged C128 using clonNAT as yeast selection marker. To screen for positive clones, the C128 genomic regions harbouring the desired point mutations were amplified by colony PCR, gel purified and analysed by Sanger sequencing (Eurofins Genomics).

4.1.8. Yeast spotting assays

Yeast spotting assays were used to assess viability of yeast Pol III mutants and were performed by Dr. Helga Grötsch. Yeast cultures were inoculated over night in YPD medium. The next day, the cultures were diluted to an optical density (OD) of 0.2. After 4 hours, the cultures were diluted to an OD of 0.2 using PBS. Subsequently, a ten-fold serial dilution (3 steps) using PBS was performed. Diluted cultures (starting from the OD of 0.02) were applied as single 8 µL drops onto YPAD plates, which were incubated at 25 °C, 30 °C and 37 °Cfor 3 days before inspecting the growth phenotypes.

4.2. Protein expression and purification

4.2.1. Expression of recombinant human Pol III

Human Pol III genes were co-expressed in human Expi293F (Thermo Fisher Scientific) suspension cells. Handling of the Expi293F cells and transfections were carried out by Dr. Helga Grötsch. Expi293F cells were grown in Expi293F expression medium (Thermo Fisher Scientific) in 125 mL disposable Erlenmeyer flasks (Corning) for maintenance or in 1 L disposable roller bottles (JetBiofil) for protein expression (250 to 330 mL cells per 1 L bottle) in a Minitron incubator shaker, set to 37 °C, 8% CO_2, 70% humidity, 125 RPM. The multi-gene expression plasmids p1769 (a 14-gene construct containing the 3x FLAG-tagged Pol III core and heterodimer genes) and p1767 (a 3-gene construct containing the Pol III heterotrimer) were used for co-transfection. Plasmid DNA was produced and purified bwith plasmid midiprep kits (Qiagen). For a 500 mL transfection, the cells were diluted to a cell density of $1x10^6$ cells/mL 24 hours prior to the transfection. Under a laminar flow hood, 225 µg of p1767 and 225 µg of p1769 were mixed in 16.7 mL pre-warmed expression medium, 1483 µL PEI (1 µ/ µL) were added. The mixture was incubated for 15 minutes at room temperature and slowly added to the cells. Protein expression was carried out for 48 hours. The cells were harvested by centrifugation (500 g, room temperature), snap-frozen in liquid nitrogen, and stored at -80 °C until usage.

4.2.2. Purification of recombinant human Pol III

Frozen Expi293F cells were thawed in a cold water bath, resuspended in cold 100 mL lysis buffer (25 mM Hepes pH 7.5, 150 mM AmSO4, 5 mM MgCl2, 5% Glycerol, 0.2% Triton-X 100, 1 cOmplete, EDTA-free Protease Inhibitor Cocktail Tablet (Merck) per 50 mL, 0.1 mM PMSF), and lysed by sonification (40% amplitude, 15 s on, 15 s off, repeat 5 times). The lysate was cleared by centrifugation (45 minutes, 60,000 g, 4 °C). The supernatant was filtered through a 1.2 µm syringe filter (GE Healthcare) and incubated for 2.5 hours

at 4 °C with 1.5 mL ANTI-FLAG M2 resin (Sigma-Aldrich, Merck), pre-equilibrated in wash buffer (25 mM Hepes pH 7.4, 150 AmSO4, 5 mM MgCl2, 5% Glycerol), while gently rolling the incubation tube. The mixture was transferred to a gravity flow column (Bio-Rad), the flow-though was removed, and the resin was washed with 100 mL wash buffer. Bound proteins were eluted by incubating the resin for 5 minutes with 1 mL of FLAG-elution buffer (25 mM Hepes pH 7.5, 150 mM AmSO4, 5 mM MgCl2, 5% Glycerol, 166 µg/mL 3x FLAG peptide (Sigma-Aldrich, Merck). The FLAG-eluate was collected in a fresh tube and the elution was repeated another two times, followed by a 2 mL resin wash with wash buffer, which was also collected. DTT was directly added to the eluted samples to a final concentration of 1 mM. Eluted samples were pooled (volume = 5 mL) and 10 mL of Mono Q A buffer (25 mM Hepes pH 7.5, 50 mM AmSO4, 5 mM MgCl2, 5% Glycerol, 1 mM DTT) were added. The sample was applied to a Mono Q 5/50 GL anion exchange column (Sigma-Aldrich, Merck), equilibrated in Mono Q A buffer, and eluted from the column using a 40 column volume (CV) linear gradient ranging from 50 mM AmSO4 to 1 M AmSO4. Eluted fractions were analysed via SDS-PAGE), and the fraction containing stoichiometric human Pol III were pooled, concentrated in an Amicon Ultra 100 KDa cutoff concentrator (Merck), buffer exchanged to 25 mM Hepes pH 7.5, 150 AmSO4, 5 mM MgCl2, 5% Glycerol, 2 mM DTT, concentrated to a volume of 50 µL, snap frozen in liquid nitrogen, and stored at -80 °C until usage.

4.2.3. Purification of endogenous human Pol III

Endogenous human Pol III for structural analysis via cryo-EM was purified by Agata D. Misiaszek from a Expi293F cell line, which was genomically engineered via CRISPR-Cas9 to carry a mCherry-Strep II-6xHis tag on the C-terminus of the Pol III subunit RPAC1 as described [162].

4.2.4. Large scale Purification of yeast Pol III

Endogenous yeast Pol III was purified in large scale from a 100 L fermenter culture as described [193]. Yeast cells were fermented by Rene Wetzel. In brief, yeast cells containing a C-terminal TAP-tag on subunit AC40 were lysed using glass beads and a bead beater (BioSpec). The lysate was cleared via centrifugation and loaded on a heparin column (Cytiva, Merck), followed by IgG affinity purification using TEV protease to elute the bound proteins from the IgG resin (Cytiva, Merck) and finally anion exchange chromatography (Mono Q 10/100 GL, Sigma-Aldrich, Merck) to separate Pol I and Pol III. The following modification were conducted: For resuspension of yeast cells cells and equilibration of the heparin column, the following buffer was used: 250 mM Tris–HCl pH

8, 20% glycerol, 250 mM AmSO4, 1 mM EDTA, 10 mM MgCl2,10 µM ZnCl2, 12 mM mercaptoethanol), supplemented with protease inhibitor cocktail (0.3 ug/mL Leupeptin, 1.4 ug/mL Pepstatin, 170 µg/mL PMSF, 330 µg/mL Benzamidin, in EtOH). After the Mono Q column, eluted Pol III was buffer exchanged into: 15 mM Hepes, pH 7.5, 150 mM AmSO4, 10 mM DTT, concentrated to a concentration of 7.5 to 14.9 mg/mL, snap-frozen in liquid nitrogen and stored at -80°C until usage.

4.2.5. Small scale Purification of yeast Pol III

Endogenous yeast Pol III was purified in small scale from 16 L yeast cultures grown by Dr. Helga Grötsch in conical flasks for structure-function studies. *S. cerevisiae* strains harbouring a C-terminal TAP tag on subunit C128 (either wild-type (WT) or mutated) were used for protein purification. Frozen cells (typically 50 g per 16 L culture) were thawed over night at 4 °C and resuspended in lysis buffer: 250 mM Tris–HCl pH 8, 20% glycerol, 250 mM AmSO4, 1 mM EDTA, 10 mM MgCl2, 10 µM ZnCl2, 12 mM upbeta-mercaptoethanol), supplemented with protease inhibitor cocktail (0.3 ug/mL Leupeptin, 1.4 ug/mL Pepstatin, 170 µg/mL PMSF, 330 µg/mL Benzamidin, in EtOH). The cells were lysed using glass beads and a bead beater, cleared by centrifugation and loaded on two 5 mL HiTrap Heparin HP (Cytiva, Merck), connected in series that was equilibrated in lysis buffer. Proteins were eluted using a 100 CV linear gradient (250 mM AmSO4 to 1 M AmSO4). Pol III was further purified as described above with the exception that the anion exchange chromatography step was omitted.

4.2.6. Purification of yeast Pol III Δ

Endogenous yeast Pol III Δ, which lacks subunits C53-C37/C11, was purified from a 100 L fermenter culture, which was maintained by Rene Wetzel. The *S. cerevisiae* Pol III Δ strain was a gift from Richard Maraia (National Institutes of Health, Bethesda, MD, USA) and harbours an N-terminal 6x His tag on subunit C128, which was used for protein purification. Frozen cells (330 g) were thawed over night at 4 °C and resuspended in lysis buffer: 60 mM Hepes pH 7.5, 750 mL NaCl, 10.5 mM MgCl2, 7.5% glycerol, 14.3 mM upbeta-mercaptoethanol, supplemented with protease inhibitor cocktail (0.3 ug/mL Leupeptin, 1.4 ug/mL Pepstatin, 170 µg/mL PMSF, 330 µg/mL Benzamidin, in EtOH). The cells were lysed using glass beads and a bead beater and cleared by two centrifugation runs: first run at 30,000 g, 1 hour, 4 °C; second run at 72,465 g, 1 hour, 4 °C. The cleared lysate was loaded multiple times over night (in a looped setup) on a 5 mL HisTrap HP column (Cytiva, Merck) equilibrated in 20 mM Hepes pH 7.5, 500 mM NaCl, 7 mM MgCl2, 10% glycerol, 10 mM upbeta-mercaptoethanol, 10 mM imidazole.

Bound proteins were eluted using a 20 CV linear (10 mM to 300 mM imidazole). Pooled fractions were diluted with 120 mL Mono Q A buffer (40 mM Tris-HCl pH 8, 68 mM AmSO4, 0.5 mM EDTA, 1 mM MgCl2, 10 µM ZnCl2, 10 mM DTT) and applied to a Mono Q 5/50 GL column. A 30 CV linear gradient (68 mM AmSO4 to 1 M AmSO4) was used to elute Pol III Δ, which was applied a second time to the Mono Q 5/50 GL column after diluting pooled fraction with Mono Q A buffer. A two-step linear gradient (fist step: 68 mM to 250 mM AmSO4; second step: 250 mM to 1 M AmSO4) was used for elution to increase purity. Pol III Δ containing fraction were pooled, buffer exchanged into: 15 mM Hepes, pH 7.5, 150 mM AmSO4, 10 mM DTT, concentrated to a concentration of 2.4 mg/mL, snap-frozen in liquid nitrogen and stored at -80°C until usage.

4.3. Biochemistry

4.3.1. Promoter-independent transcription assays

Promoter-independent transcription assays on tailed-template or single-stranded DNA constructs to test the activity of recombinant and endogenous human Pol III were carried out by Dr. Florence Baudin. Tailed-template DNA constructs were designed to carry a single-stranded DNA overhang. 2 pmol of DNA, dissolved in H2O, were incubated with 2 to 3 pmol of Pol III, incubated at room temperature for 10 minutes, and 5 µL of transcription start buffer (20 mM Hepes pH 7.5, 10 mM MgSO4, 10% Glycerol, 10 mM DTT), 1 µL of NTP mix (2 mM ATP, 2 mM CTP, 2 mM GTP, 0.1 mM UTP), and 1 µL of [α-P32]UTP (10 mCi/ml, Hartmann Analytic) were added. The reaction was incubated for 30 minutes at 28 or 37 °C, stopped by adding 10 uL of formamide loading buffer, boiled for 3 minutes at 95 °C, and loaded on a 10% denaturing PAGE. For promoter-independent transcription assays on yeast Pol III, I also used tailed template DNA constructs. 1 µL of template and non-template strands (30 µM) were mixed, boiled for 2 minutes at 95 °C, placed on ice for 1 minute, 2 µL of hybridization buffer (40 mM Hepes pH 7.5, 24 mM MgCl2, 200 mM NaCl, 20 mM DTT) were added, and the mixture was placed at room temperature for 10 minutes. Afterwards, 11 µL of transcription reaction buffer (TRB) (20 mM Hepes, pH 7.6, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 10 mM DTT) were added and the annealed DNA was kept on ice until usage. Next, 2 pmol of DNA were mixed with 2.5 pmol WT or mutant Pol III in a 5 µL reaction volume and incubated at room temperature for 10 minutes. RNA synthesis was initiated by adding 3.9 µL of TRB, 0.7 µL of NTP mix (10 mM ATP, 10 mM CTP, 1 mM UTP) and 0.7 µL of [α-P32]UTP (10 mCi/ml). The reaction was incubated at room temperature for 40 minutes and stopped by adding 138 µL of stop buffer (0.6 M AcNa, 30 mM EDTA, 0.5% EDTA), followed by phenol–chloroform extraction, boiling at

95 °C for 3 minutes and loading on a 15% denaturing PAGE. For all assays, radioactivity was captured on a phosphor-imaging screen (Fujifilm) and detected with a Typhon FLA9500 (GE Healthcare). Band intensities were quantified using Image Lab (Bio-Rad) and plotted with the statistical computing software R. For normalization, band intensities were divided by the total intensities of the lanes containing the respective bands and multiplied by a factor of 1000 for illustration purposes. Tailed-template experiments on yeast Pol III were performed in triplicates.

4.3.2. Promoter-dependent transcription assays

For promoter-dependent transcription assays, a double-stranded (ds) DNA hybrid construct was used containing the *S. cerevisiae* U6 snRNA promoter region and downstream of it a portion of the *S. cerevisiae* tDNA gene tL(CAA)L. Nucleic acid strand annealing was performed as described above. Yeast Pol III was diluted to 5 µM in TRB (20 mM Hepes, pH 7.6, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 10 mM DTT). For recruitment of Pol III to the promoter DNA, yeast TFIIIB was freshly assembled using Bdp1 and Brf1-TBP (kindly provided and purified by Dr. Matthias K. Vorländer as described [77]). For TFIIIB assembly, 2.4 µL Bdp1 (6 mg/mL), 2.8 µL Brf1-TBP (6 mg/mL) and 4.8 µL TRB-300 buffer (20 mM Hepes, pH 7.6, 300 mM NaCl, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 10 mM DTT) were mixed, incubated on ice for 10 minutes and diluted 4-fold in TRB-300 to 5 µM. Next, 2.5 pmol of dsDNA were incubated with 5 pmol of TFIIIB, incubated for 10 minutes on ice, followed by the addition of 5 pmol Pol III and incubation on ice for at least 10 minutes. Transcription was initiated by adding 5.4 µL of TRB, 0.7 µL of NTP mix (10 mM ATP, 10 mM CTP, 1 mM UTP) and 0.7 µL of [α-P32]UTP (10 mCi/ml), and the reaction was incubated at room temperature or 28 °C. Typically, time courses were conducted, for which individual reactions were stopped at indicated time points by adding 4 µL formamide loading buffer and boiling for 3 min at 95 °C, followed by loading on a 15% denaturing PAGE. Radioactivity was captured on a phosphor-imaging screen and detected with a Typhon FLA9500.

4.3.3. Transcription scaffold preparation

For structural analysis of nucleic acid-bound Pol III (both human and yeast), transcription scaffolds were freshly prepared in a thermocycler (Bio-Rad) using HPLC-grade nucleic acid oligonucleotides (Sigma-Aldrich). 1 µL Template- and 1 µL nontemplate DNA (100 µM each) were incubated at a temperature of 95 °C, which was incrementally decreased to 25 °C at a rate of 1 °C per minute. 3 µl of hybridization buffer (see above) and 1 µl of pre-heated RNA (100 µM) were added, the mixture was incubated at 45 °C for 3 min,

cooled down to 20 °C at a rate of 0.7 °C per minute and the transcription scaffold was kept on ice until usage.

4.3.4. RNA extension and cleavage assays

RNA extension assays were used to analyse the activity of human and yeast Pol III. Extension assays on human Pol III were carried out by Dr. Florence Baudin. RNA primers were labeled with [γ-32P]ATP (10 mCi/ml, Hartmann Analytic) and T4 PNK (NEB) and purified via PAGE. 1 µL of diluted radioactively-labeled RNA (equalling 50-100 counts per second) were mixed with 1 µL of template and 1 µL non-template DNA (20 µM, each), boiled for 5 minutes at 95 °C, placed on ice for 1 minute, and 3 µL of hybridization buffer (see above) were added. Typically, the scaffold was diluted to 2 µM using transcription reaction buffer (20 mM Hepes, pH 7.6, 60 mM AmSO4, 10 mM MgSO4, 10% Glycerol, 2 mM DTT), and 1 µL was mixed with 1 µL of Pol III (4 µM). After 10 minutes of incubation at room temperature, 6 µL of transcription start buffer (TRB supplemented with NTPs (1.3 mM each) was added, and the reaction was incubated at 28 °C for 25 minutes. The reaction was stopped by adding 4 µL of formamide loading buffer, boiled for 3 minutes at 95 °C and loaded on a 17% denaturing PAGE. Radioactivity was captured on a phosphor-imaging screen and detected with a Typhon FLA9500.

4.4. Structure determination via electron microscopy

4.4.1. Negative-stain electron microscopy

Negative-stain electron microscopy (EM) on human and yeast Pol III was performed to assess sample homogeneity. Samples were diluted to 50 nM in EM buffer (15 mM Hepes pH 7.5, 150 mM AmSO4, 5 mM MgCl2, 10 mM DTT), and 3.5 µL were applied to a carbon-coated 400 mesh copper grid (Plano), which was glow discharged using a Pelco EasyGlow instrument (Ted Pella, Inc). After 60 s incubation, the grids were washed with EM buffer, stained with 20 µL of 1% (w/v) uranyl acetate and air dried. Negative-stain EM images were recorded with a 4 × 4,000 CCD camera on a Tecnai T12 transmission electron microscope (TEM) operated at 120 keV (Thermo Fisher Scientific, FEI) using the following parameters: Defocus: 1 µm, magnification: 68,000×, pixel size: 1.6 Å per pixel, dose rate: 40 electrons per $Å^2$.

Table 4. Cryo-EM sample preparation conditions. Hs - *H. sapiens*; Sc - *S. cerevisiae*; EC - Elongation complex; PTC - Pre-Termination Complex, Δ - Δ C53-C37/C11

	HsPol III Apo	**HsPol III Apo**	**HsPol III EC**	
Grid type	Au R1.2/1.3 grid	Cu R2/1 grid	Cu R2/1 grid	
Meshsize	300	200	200	
Manufacturer	Quantifoil	Quantifoil	Quantifoil	
Grid treatment	Plasma-cleaned	Plasma-cleaned	Plasma-cleaned	
Concentration	Pol III: 1.6 uM	Pol III: 1.6 uM	Pol III: 2.7 uM Scaffold: 2.7 uM	
	ScPol III PTC	**ScPol III EC**	**ScPol III + NTPs**	**ScPol IIIΔ**
Grid type	Cu R2/1 grid	Cu R2/1 grid	Cu R2/1 grid	Cu R2/1 grid
Meshsize	200	200	200	200
Manufacturer	Quantifoil	Quantifoil	Quantifoil	Quantifoil
Grid treatment	Glow-discharged	Glow-discharged	Glow-discharged	Glow-discharged
Concentration	Pol III: 2.5 uM Scaffold: 2.8 uM	Pol III: 2.8 uM Scaffold: 3 uM	Pol III: 3.4 uM Scaffold: 3.6 uM	Pol III: 3.4 uM Scaffold: 3.6 uM

4.4.2. Cryo-EM sample preparation

Human and yeast Pol III cryo-EM samples were prepared on a Vitrobot Mark IV (Thermo Fisher Scientific, FEI) using the following parameters: 100% humidity; 4°C, 10 seconds wait time, 4 seconds blot time, blot force 4. All cryo-EM samples were in 15 mM Hepes, pH 7.5, 150 mM NH4SO4, 10 mM DTT, 5 mM MgCl2 supplemented with 4 mM CHAPSO (final concentration) and plunge-frozen in liquid ethane. For nucleic acid bound complexes, Pol III samples were mixed with transcription scaffolds prepared as described above. Sample concentrations and grid types are listed in Table 4.

4.4.3. Cryo-EM data collection and processing

Cryo-EM data collection parameters are summarized in Table 5. Automated data acquisition was performed using SerialEM [194]. Micrographs were pre-processed (initial frame alignment, contrast transfer function (CTF)-estimation, doseweighting, automated particle picking) was performed using Warp [165]. Images were further processed in RELION 3.1 [166] by importing the particle stacks generated by Warp. Micrographs were also frame-aligned, dose-weighted and CTF-corrected in RELION to run particle polishing a later stage during processing. Cryo-EM processing pipelines typically involved multiple rounds of unmasked and masked 3D classifications runs. For particles that went into the final 3D reconstructions, two iterative rounds of particle polishing followed by CTF refinement [195] were performed. To improve map qualities of flexible regions, the final 3D reconstructions were subjected to RELION's multi-body refinement [167]. Soft

masks for masked 3D classifications and multi-body refinement were created with the molmap command in UCSF Chimera [196] and the mask creation option in RELION. Nominal map resolution values were calculated using RELION post-processing, which is based on Fourier shell correlation (FSC) using the 0.143 cut-off criterion [197] and global B-factor sharpening to reduce contrast loss at higher resolutions. Local resolution ranges were calculated using RELION or MonoRes [198], integrated into the Scipion software framework [199], which was run using SBGrid software collection [200]. The cryoSPARC [164] software platform was also used for cryo-EM processing, in particular for initial assessment of dataset quality via 2D classification, 3D ab-initio reconstruction generations, non-uniform refinement [201], and 3D variability analysis [168].

Table 5. Cryo-EM data collection parameters. Hs - *H. sapiens*; Sc - *S. cerevisiae*; EC - Elongation complex; PTC - Pre-Termination Complex, Δ - Δ C53-C37/C11

	HsPol III Apo	**HsPol III Apo**	**HsPol III EC**	
Microscope	Talos arctica	Titan Krios	Titan Krios	
Detector	Falcon III	Gatan K3	Gatan K2	
Collection mode	Linear	Counting	Counting	
Magnification	92,000x	105,000x	130,000x	
Voltage (kV)	200	300	300	
Defocus range (μm)	1.00-3.50	0.75-2.25	0.75-2.00	
Pixel size (Å)	1.56	0.82	1.05	
Micrographs (no.)	753	9,544	9,488	
Frames (no.)	12	40	40	
Dose e/Å2/frame	3.91	0.95	1.01	
	ScPol III PTC	**ScPol III EC**	**ScPol III + NTPs**	**ScPol IIIΔ**
Microscope	Titan Krios	Titan Krios	Titan Krios	Titan Krios
Detector	Gatan K2	Gatan K2	Gatan K2	Gatan K3
Collection mode	Counting	Counting	Counting	Counting
Magnification	130,000x	130,000x	130,000x	81,000x
Voltage (kV)	300	300	300	300
Defocus range (μm)	0.75 - 2.25	0.50-1.75	0.75-2.25	1.00-2.50
Pixel size (Å)	1.04	1.04	1.04	1.05
Micrographs (no.)	8,446	8,700	11,840	11,833
Frames (no.)	29	30	40	40
Dose e/Å2/frame	1.41	1.27	1.12	0.99

4.4.4. Structural model building and refinement

For model building, cryo-EM maps were sharpened with LocalDeblur [202], implemented in Scipion, or with density modification [175], implemented into Phenix [203].
For human Pol III, an initial model was generated using available structures (RPC3-RPC7 (PDB: 5AFQ) and of RPABC1, RPABC2, RPABC3, RPABC4, RPABC5 (PDB: 5IY6)), homology models obtained from the SWISSMODEL Repository [204] (RPC1, RPC2,

RPC4, RPAC2) and homology models generated with Phyre2 [205] (RPC5, RPC6, RPC7, RPC8, RPC9, RPC10, RPAC1). The protein models were aligned to the respective chains of the yeast Pol III Apo structure (PDB: 5FJ9) and individually fitted in the cryo-EM map as rigid bodies using Chimera. Next, COOT [206, 207] was used for model building and real-space refinement. For the human Pol III EC, the DNA:RNA-duplex of the yeast Pol III EC structure (PDB: 5FJ8) was placed into the human Pol III EC map and automatically fitted and refined using interatomic distance restraints generated with ProSMART [208]. The nucleic acid sequence was mutated to match the sequence of used transcription scaffold, followed by manual fitting.
For subunit RPC10, two conformations of the C-terminal Zn-Ribbon domain (C-ribbon) were observed (inside and outside the Pol III funnel domain). Consequently, two models of the human Pol III EC, harbouring different conformations of the C-ribbon domain, were build by manually placing a homology model into the densities, followed by building and refinement in COOT.
The atomic model of the human heterotrimer could be extended by building the FeS domain of RPC6, and the stalk bridge and core binding region of RPC7. A cubane 4Fe-4S cluster (PDB: 2B3Y) was placed into a strong density signal. The RPC6 FeS domain and the RPC7 extensions were build *de novo*.
For the model of the human heterodimer, subunit RPC5 could be C-terminally extended by building the first two winged helix domains (WH) into a map that was derived via masked 3D classification using generated homology models as a starting point. The build WH1-WH2 domain was also used to obtain human Pol III harbouring an alternative conformation of the RPC5 C-terminal extension, which was also derived via masked 3D classification. Additional density close to Pol III jaw domain and the heterodimer dimerization module (DM) could be assigned to an N-terminal extension of RPC4 (termed DM*) and was build *de novo*.
For building of the yeast Pol III PTC, Pol III EC (control), Pol III PTC plus NTPs and Pol III Δ, the available structure of the yeast Pol III EC (PDB: 5FJ8) was used as a starting model. In case of the Pol III PTC, 7 bases of the unwound non-template (NT) strand could be build. The upstream DNA was build using a map of the Pol III PTC plus NTPs that showed significant improvement of the upstream DNA signal after masked classification. To improve the clash scores, build models were subjected to ISOLDE [209], which is based on flexible fitting of proteins and nucleic acids using molecular-dynamics and implemented into ChimeraX [210], followed by real-space refinement in Phenix.

References

1. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nature Reviews Microbiology* **9,** 85–98 (2011).
2. Choi, K. H. Viral polymerases. *Advances in Experimental Medicine and Biology* **726,** 267–304 (2012).
3. Zong, J., Yao, X., Yin, J., Zhang, D. & Ma, H. Evolution of the RNA-dependent RNA polymerase (RdRP) genes: Duplications and possible losses before and after the divergence of major eukaryotic groups. *Gene* **447,** 29–39 (2009).
4. Börner, T., Aleynikova, A. Y., Zubo, Y. O. & Kusnetsov, V. V. Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847,** 761–769 (2015).
5. Griesenbeck, J., Tschochner, H. & Grohmann, D. Structure and Function of RNA Polymerases and the Transcription Machineries. *Subcellular Biochemistry* **83,** 225–270 (2017).
6. Chen, F. X., Smith, E. R. & Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. *Nature Reviews Molecular Cell Biology* **19,** 464–478 (2018).
7. Nudler, E. RNA Polymerase Backtracking in Gene Regulation and Genome Instability. *Cell* **149,** 1438–1445 (2012).
8. Zhou, M. & Law, J. A. RNA Pol IV and V in gene silencing: Rebel polymerases evolving away from Pol II's rules. *Current Opinion in Plant Biology* **27,** 154–164 (2015).
9. Cramer, P. *et al.* Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288,** 640–649 (2000).
10. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292,** 1876–1882 (2001).
11. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. *Science* **292,** 1863–1876 (2001).
12. Plaschka, C. *et al.* Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518,** 376–380 (2015).
13. Plaschka, C. *et al.* Transcription initiation complex structures elucidate DNA opening. *Nature* **533,** 353–358 (2016).
14. He, Y. *et al.* Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533,** 359–365 (2016).
15. Engel, C., Plitzko, J. & Cramer, P. RNA polymerase I–Rrn3 complex at 4.8 Å resolution. *Nature Communications* **7,** 12129 (2016).
16. Hoffmann, N. A. *et al.* Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528,** 231–236 (2015).
17. Fernández-Tornero, C. *et al.* Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502,** 644–649 (2013).

18. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural Basis of Transcription: Role of the Trigger Loop in Substrate Specificity and Catalysis. *Cell* **127,** 941–954 (2006).
19. Kettenberger, H., Armache, K.-J. & Cramer, P. Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS. *Molecular Cell* **16,** 955–965 (2004).
20. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224,** 234–237 (1969).
21. Sklar, V. E., Schwartz, L. B. & Roeder, R. G. Distinct molecular structures of nuclear class I, II, and III DNA-dependent RNA polymerases. *Proceedings of the National Academy of Sciences* **72,** 348–352 (1975).
22. Hager, G. L., Holland, M. J. & Rutter, W. J. Isolation of ribonucleic acid polymerases I, II, and III from *Saccharomyces cerevisiae*. *Biochemistry* **16,** 1–8 (1977).
23. Weinmann, R. & Roeder, R. G. Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. *Proceedings of the National Academy of Sciences* **71,** 1790–1794 (1974).
24. Weinmann, R., Raskas, H. J. & Roeder, R. G. Role of DNA-dependent RNA polymerases II and III in transcription of the adenovirus genome late in productive infection. *Proceedings of the National Academy of Sciences* **71,** 3426–3439 (1974).
25. Brown, D. D., Wensink, P. C. & Jordan, E. Purification and some characteristics of 5S DNA from *Xenopus laevis*. *Proceedings of the National Academy of Sciences* **68,** 3175–3179 (1971).
26. Nath, K. & Bollon, A. P. Organization of the yeast ribosomal RNA gene cluster via cloning and restriction analysis. *Journal of Biological Chemistry* **252,** 6562–6571 (1977).
27. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Research* **44,** D184–D189 (2016).
28. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136,** 701–718 (2009).
29. Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proceedings of the National Academy of Sciences* **83,** 8575–8579 (1986).
30. Söderlund, H., Pettersson, U., Vennström, B., Philipson, L. & Mathews, M. B. A new species of virus-coded low molecular weight RNA from cells infected with adenovirus type 2. *Cell* **7,** 585–593 (1976).
31. Thimmappaya, B., Weinberger, C., Schneider, R. J. & Shenk, T. Adenovirus VAI RNA is required for efficient translation of viral mRNAs at late times after infection. *Cell* **31,** 543–551 (1982).
32. Lerner, M. R., Andrews, N. C., Miller, G. & Steitz, J. A. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proceedings of the National Academy of Sciences* **78,** 805–809 (1981).

33. Howe, J. & Shu, M.-D. Epstein-Barr virus small RNA (EBER) genes: Unique transcription units that combine RNA polymerase II and III promoter elements. *Cell* **57,** 825–834 (1989).
34. Iwakiri, D. & Takada, K. Role of EBERs in the pathogenesis of EBV infection. *Advances in Cancer Research* **107,** 119–136 (2010).
35. Walter, P. & Blobel, G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299,** 691–698 (1982).
36. Briand, J.-F., Navarro, F., Gadal, O. & Thuriaux, P. Cross Talk between tRNA and rRNA Synthesis in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **21,** 189–195 (2001).
37. Bartkiewicz, M., Gold, H. & Altman, S. Identification and characterization of an RNA molecule that copurifies with RNase P activity from HeLa cells. *Genes & Development* **3,** 488–499 (1989).
38. Yuan, Y. & Reddy, R. 5' flanking sequences of human MRP/7-2 RNA gene are required and sufficient for the transcription by RNA polymerase III. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1089,** 33–39 (1991).
39. Rome, L., Kedersha, N. & Chugani, D. Unlocking vaults: organelles in search of a function. *Trends in Cell Biology* **1,** 47–50 (1991).
40. Horos, R. *et al.* The Small Non-coding Vault RNA1-1 Acts as a Riboregulator of Autophagy. *Cell* **176,** 1054–1067.e12 (2019).
41. Wolin, S. L. & Steitz, J. A. Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell* **32,** 735–744 (1983).
42. Boccitto, M. & Wolin, S. L. Ro60 and Y RNAs: structure, functions, and roles in autoimmunity. *Critical Reviews in Biochemistry and Molecular Biology* **54,** 133–152 (2019).
43. Stein, A. J., Fuchs, G., Fu, C., Wolin, S. L. & Reinisch, K. M. Structural Insights into RNA Quality Control: The Ro Autoantigen Binds Misfolded RNAs via Its Central Cavity. *Cell* **121,** 529–539 (2005).
44. Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T. & Schmid, C. W. Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. *Nature* **284,** 372–374 (1980).
45. Pan, J., Elder, J. T., Duncan, C. H. & Weissman, S. M. Structural analysis of interspersed repetitive polymerase III transcription units in human DNA. *Nucleic Acids Research* **9,** 1151–1170 (1981).
46. Roy-Engel, A. M. *et al.* Alu insertion polymorphisms for the study of human genomic diversity. *Genetics* **159,** 279–290 (2001).
47. Dewannieux, M., Esnault, C. & Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nature Genetics* **35,** 41–48 (2003).
48. Sutcliffe, J. G., Milner, R. J., Gottesfeld, J. M. & Lerner, R. A. Identifier sequences are transcribed specifically in brain. *Nature* **308,** 237–241 (1984).
49. Lin, D., Pestova, T. V., Hellen, C. U. T. & Tiedge, H. Translational control by a small RNA: dendritic BC1 RNA targets the eukaryotic initiation factor 4A helicase mechanism. *Molecular and Cellular Biology* **28,** 3008–3019 (2008).
50. Shin, H., Kim, Y., Kim, M. & Lee, Y. BC200 RNA: An Emerging Therapeutic Target and Diagnostic Marker for Human Cancer. *Molecules and Cells* **41,** 993–999 (2018).

51. Dergai, O. & Hernandez, N. How to Recruit the Correct RNA Polymerase? Lessons from snRNA Genes. *Trends in Genetics* **35,** 457–469 (2019).
52. Arimbasseri, A. G. & Maraia, R. J. RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends in Biochemical Sciences* **41,** 546–559 (2016).
53. Pieler, T., Oei, S. L., Hamm, J., Engelke, U. & Erdmann, V. A. Functional domains of the *Xenopus laevis* 5S gene promoter. *The EMBO Journal* **4,** 3751–3756 (1985).
54. Pieler, T., Hamm, J. & Roeder, R. G. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* **48,** 91–100 (1987).
55. Engelke, D. R., Ng, S. Y., Shastry, B. S. & Roeder, R. G. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* **19,** 717–728 (1980).
56. Galli, G., Hofstetter, H. & Birnstiel, M. L. Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. *Nature* **294,** 626–631 (1981).
57. Palida, F. A., Hale, C. & Sprague, K. U. Transcription of a silkworm tRNA(cAla) gene is directed by two AT-rich upstream sequence elements. *Nucleic Acids Research* **21,** 5875–5881 (1993).
58. Huibregtse, J. M., Evans, C. F. & Engelke, D. R. Comparison of tRNA gene transcription complexes formed in vitro and in nuclei. *Molecular and Cellular Biology* **7,** 3212–3220 (1987).
59. Chalker, D. L. & Sandmeyer, S. B. Sites of RNA polymerase III transcription initiation and Ty3 integration at the U6 gene are positioned by the TATA box. *Proceedings of the National Academy of Sciences* **90,** 4927–4931 (1993).
60. Eschenlauer, J. B., Kaiser, M. W., Gerlach, V. L. & Brow, D. A. Architecture of a yeast U6 RNA gene promoter. *Molecular and Cellular Biology* **13,** 3015–3026 (1993).
61. Krol, A., Carbon, P., Ebel, J. P. & Appel, B. *Xenopus tropicalis* U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes. *Nucleic Acids Research* **15,** 2463–2478 (1987).
62. Sadowski, C. L., Henry, R. W., Lobo, S. M. & Hernandez, N. Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. *Genes & Development* **7,** 1535–1548 (1993).
63. Henry, R. W., Mittal, V., Ma, B., Kobayashi, R. & Hernandez, N. SNAP19 mediates the assembly of a functional core promoter complex (SNAPc) shared by RNA polymerases II and III. *Genes & Development* **12,** 2664–2672 (1998).
64. Kunkel, G. R. & Pederson, T. Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used. *Genes & Development* **2,** 196–204 (1988).
65. Dergai, O. *et al.* Mechanism of selective recruitment of RNA polymerases II and III to snRNA gene promoters. *Genes & Development* **32,** 711–722 (2018).
66. Schramm, L., Pendergrast, P. S., Sun, Y. & Hernandez, N. Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes & Development* **14,** 2650–2663 (2000).

67. Hoffmann, N. A., Jakobi, A. J., Vorländer, M. K., Sachse, C. & Müller, C. W. Transcribing RNA polymerase III observed by electron cryomicroscopy. *The FEBS Journal* **283,** 2811–2819 (2016).
68. Jasiak, A. J., Armache, K.-J., Martens, B., Jansen, R.-P. & Cramer, P. Structural Biology of RNA Polymerase III: Subcomplex C17/25 X-Ray Structure and 11 Subunit Enzyme Model. *Molecular Cell* **23,** 71–81 (2006).
69. Landrieux, E. *et al.* A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *The EMBO Journal* **25,** 118–128 (2006).
70. Wang, Z. & Roeder, R. G. Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. *Genes & Development* **11,** 1315–1326 (1997).
71. Kuhn, C.-D. *et al.* Functional Architecture of RNA Polymerase I. *Cell* **131,** 1260–1272 (2007).
72. Geiger, S. R. *et al.* RNA Polymerase I Contains a TFIIF-Related DNA-Binding Subcomplex. *Molecular Cell* **39,** 583–594 (2010).
73. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Molecular Cell* **45,** 439–446 (2012).
74. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nature Reviews Molecular Cell Biology* **16,** 129–143 (2015).
75. Kassavetis, G. A., Prakash, P. & Shim, E. The C53/C37 Subcomplex of RNA Polymerase III Lies Near the Active Site and Participates in Promoter Opening. *Journal of Biological Chemistry* **285,** 2695–2706 (2010).
76. Wu, C.-C., Lin, Y.-C. & Chen, H.-T. The TFIIF-Like Rpc37/53 Dimer Lies at the Center of a Protein Network To Connect TFIIIC, Bdp1, and the RNA Polymerase III Active Center. *Molecular and Cellular Biology* **31,** 2715–2728 (2011).
77. Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553,** 295–300 (2018).
78. Lee, J., Moir, R. D., McIntosh, K. B. & Willis, I. M. TOR Signaling Regulates Ribosome and tRNA Synthesis via LAMMER/Clk and GSK-3 Family Kinases. *Molecular Cell* **45,** 836–843 (2012).
79. Wang, Z., Wu, C., Aslanian, A., Yates, J. R. & Hunter, T. Defective RNA polymerase III is negatively regulated by the SUMO-Ubiquitin-Cdc48 pathway. *eLife* **7,** e35447 (2018).
80. Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553,** 301–306 (2018).
81. Lefèvre, S. *et al.* Structure-function analysis of hRPC62 provides insights into RNA polymerase III transcription initiation. *Nature Structural & Molecular Biology* **18,** 352–358 (2011).
82. Brun, I., Sentenac, A. & Werner, M. Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *The EMBO Journal* **16,** 5730–4741 (1997).
83. Wu, C.-C. *et al.* RNA polymerase III subunit architecture and implications for open promoter complex formation. *Proceedings of the National Academy of Sciences* **109,** 19232–19237 (2012).

84. Thuillier, V., Stettler, S., Sentenac, A., Thuriaux, P. & Werner, M. A mutation in the C31 subunit of *Saccharomyces cerevisiae* RNA polymerase III affects transcription initiation. *The EMBO journal* **14,** 351–359 (1995).
85. Chedin, S., Riva, M., Schultz, P., Sentenac, A. & Carles, C. The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. *Genes & Development* **12,** 3857–3871 (1998).
86. Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D. & Cramer, P. Evolution of two modes of intrinsic RNA polymerase transcript cleavage. *Journal of Biological Chemistry* **286,** 18701–18707 (2011).
87. Whitehall, S. K., Bardeleben, C. & Kassavetis, G. A. Hydrolytic cleavage of nascent RNA in RNA polymerase III ternary transcription complexes. *Journal of Biological Chemistry* **269,** 2299–2306 (1994).
88. Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502,** 650–655 (2013).
89. Cheung, A. C. M. & Cramer, P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471,** 249–253 (2011).
90. Tafur, L. *et al.* Molecular Structures of Transcribing RNA Polymerase I. *Molecular Cell* **64,** 1135–1143 (2016).
91. Arimbasseri, A. G. & Maraia, R. J. Distinguishing Core and Holoenzyme Mechanisms of Transcription Termination by RNA Polymerase III. *Molecular and Cellular Biology* **33,** 1571–1581 (2013).
92. Arimbasseri, A. G. & Maraia, R. J. Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Molecular Cell* **58,** 1124–1132 (2015).
93. Mishra, S. & Maraia, R. J. RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3 end dynamics during transcription termination. *Nucleic Acids Research* **47,** 310–327 (2019).
94. Verger, A., Monté, D. & Villeret, V. Take Your PIC. *Trends in Biochemical Sciences* (2021).
95. Willis, I. M. & Moir, R. D. Signaling to and from the RNA Polymerase III Transcription and Processing Machinery. *Annual Review of Biochemistry* **87,** 75–100 (2018).
96. Upadhya, R., Lee, J. & Willis, I. M. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Molecular Cell* **10,** 1489–1494 (2002).
97. Moir, R. D. *et al.* Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. *Proceedings of the National Academy of Sciences* **103,** 15044–15049 (2006).
98. Oficjalska-Pham, D. *et al.* General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. *Molecular Cell* **22,** 623–632 (2006).
99. Oler, A. J. & Cairns, B. R. PP4 dephosphorylates Maf1 to couple multiple stress conditions to RNA polymerase III repression. *The EMBO Journal* **31,** 1440–1452 (2012).

100. Vorländer, M. K. *et al.* Structural basis for RNA polymerase III transcription repression by Maf1. *Nature Structural & Molecular Biology* **27,** 229–232 (2020).
101. Cairns, C. A. & White, R. J. p53 is a general repressor of RNA polymerase III transcription. *The EMBO Journal* **17,** 3112–3123 (1998).
102. White, R. J., Trouche, D., Martin, K., Jackson, S. P. & Kouzarides, T. Repression of RNA polymerase III transcription by the retinoblastoma protein. *Nature* **382,** 88–90 (1996).
103. Crighton, D. *et al.* p53 represses RNA polymerase III transcription by targeting TBP and inhibiting promoter occupancy by TFIIIB. *The EMBO Journal* **22,** 2810–2820 (2003).
104. Chu, W. M., Wang, Z., Roeder, R. G. & Schmid, C. W. RNA polymerase III transcription repressed by Rb through its interactions with TFIIIB and TFIIIC2. *Journal of Biological Chemistry* **272,** 14755–14761 (1997).
105. Gomez-Roman, N., Grandori, C., Eisenman, R. N. & White, R. J. Direct activation of RNA polymerase III transcription by c-Myc. *Nature* **421,** 290–294 (2003).
106. Felton-Edkins, Z. A. *et al.* The mitogen-activated protein (MAP) kinase ERK induces tRNA synthesis by phosphorylating TFIIIB. *The EMBO Journal* **22,** 2422–2432 (2003).
107. Yeganeh, M. & Hernandez, N. RNA polymerase III transcription as a disease factor. *Genes & Development* **34,** 865–882 (2020).
108. Winter, A. G. *et al.* RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proceedings of the National Academy of Sciences* **97,** 12619–12624 (2000).
109. Zhong, Q. *et al.* The significance of Brf1 overexpression in human hepatocellular carcinoma. *Oncotarget* **7,** 6243–6254 (2016).
110. Khattar, E. *et al.* Telomerase reverse transcriptase promotes cancer cell proliferation by augmenting tRNA expression. *The Journal of Clinical Investigation* **126,** 4045–4060 (2016).
111. Lei, J., Chen, S. & Zhong, S. Abnormal expression of TFIIIB subunits and RNA Pol III genes is associated with hepatocellular carcinoma. *Liver Research* **1,** 112–120 (2017).
112. Johnson, S. A., Dubeau, L. & Johnson, D. L. Enhanced RNA Polymerase III-dependent Transcription Is Required for Oncogenic Transformation. *Journal of Biological Chemistry* **283,** 19184–19191 (2008).
113. Durrieu-Gaillard, S. *et al.* Regulation of RNA polymerase III transcription during transformation of human IMR90 fibroblasts with defined genetic elements. *Cell Cycle* **17,** 605–615 (2018).
114. Haurie, V. *et al.* Two isoforms of human RNA polymerase III with specific functions in cell growth and transformation. *Proceedings of the National Academy of Sciences* **107,** 4176–4181 (2010).
115. Filer, D. *et al.* RNA polymerase III limits longevity downstream of TORC1. *Nature* **552,** 263–267 (2017).
116. Bernard, G. *et al.* Mutations of POLR3A encoding a catalytic subunit of RNA polymerase Pol III cause a recessive hypomyelinating leukodystrophy. *American Journal of Human Genetics* **89,** 415–423 (2011).
117. Tétreault, M. *et al.* Recessive mutations in POLR3B, encoding the second largest subunit of Pol III, cause a rare hypomyelinating leukodystrophy. *American Journal of Human Genetics* **89,** 652–655 (2011).

118. Thiffault, I. *et al.* Recessive mutations in POLR1C cause a leukodystrophy by impairing biogenesis of RNA polymerase III. *Nature Communications* **6,** 7623 (2015).
119. Dorboz, I. *et al.* Mutation in POLR3K causes hypomyelinating leukodystrophy and abnormal ribosomal RNA regulation. *Neurology Genetics* **4,** e289 (2018).
120. Paolacci, S. *et al.* Specific combinations of biallelic POLR3A variants cause Wiedemann-Rautenstrauch syndrome. *Journal of Medical Genetics* **55,** 837–846 (2018).
121. Dauwerse, J. G. *et al.* Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nature Genetics* **43,** 20–22 (2011).
122. Ablasser, A. *et al.* RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nature Immunology* **10,** 1065–1072 (2009).
123. Chiu, Y.-H., Macmillan, J. B. & Chen, Z. J. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138,** 576–591 (2009).
124. Ogunjimi, B. *et al.* Inborn errors in RNA polymerase III underlie severe varicella zoster virus infections. *The Journal of Clinical Investigation* **127,** 3543–3556 (2017).
125. Carter-Timofte, M. E., Hansen, A. F., Christiansen, M., Paludan, S. R. & Mogensen, T. H. Mutations in RNA Polymerase III genes and defective DNA sensing in adults with varicella-zoster virus CNS infection. *Genes and Immunity* **20,** 214–223 (2019).
126. Han, Y., Yan, C., Fishbain, S., Ivanov, I. & He, Y. Structural visualization of RNA polymerase III transcription machineries. *Cell Discovery* **4,** 40 (2018).
127. Boissier, F., Dumay-Odelot, H., Teichmann, M. & Fribourg, S. Structural analysis of human RPC32β-RPC62 complex. *Journal of Structural Biology* **192,** 313–319 (2015).
128. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
129. Hu, P. *et al.* Characterization of human RNA polymerase III identifies orthologues for *Saccharomyces cerevisiae* RNA polymerase III subunits. *Molecular and Cellular Biology* **22,** 8044–8055 (2002).
130. Blombach, F. *et al.* Archaeal TFEα/β is a hybrid of TFIIE and the RNA polymerase III subcomplex hRPC62/39. *eLife* **4,** e08378 (2015).
131. Wong, R. C.-B. *et al.* A novel role for an RNA polymerase III subunit POLR3G in regulating pluripotency in human embryonic stem cells. *Stem Cells* **29,** 1517–1527 (2011).
132. Renaud, M. *et al.* Gene duplication and neofunctionalization: POLR3G and POLR3GL. *Genome Research* **24,** 37–51 (2014).
133. Petrie, J. L. *et al.* Effects on prostate cancer cells of targeting RNA polymerase III. *Nucleic Acids Research* **47,** 3937–3956 (2019).
134. Porrua, O. & Libri, D. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nature Reviews Molecular Cell Biology* **16,** 190–202 (2015).
135. Dieci, G. & Sentenac, A. Facilitated Recycling Pathway for RNA Polymerase III. *Cell* **84,** 245–252 (1996).
136. Bogenhagen, D. F. & Brown, D. D. Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell* **24,** 261–270 (1981).

137. Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D. & Bogenhagen, D. F. Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell* **34,** 829–835 (1983).
138. Watson, J. B., Chandler, D. W. & Gralla, J. D. Specific termination of in vitro transcription by calf thymus RNA polymerase III. *Nucleic Acids Research* **12,** 5369–5384 (1984).
139. Allison, D. S. & Hall, B. D. Effects of alterations in the 3' flanking sequence on in vivo and in vitro expression of the yeast SUP4-o tRNATyr gene. *The EMBO Journal* **4,** 2657–2664 (1985).
140. James, P. & Hall, B. D. ret1-1, a yeast mutant affecting transcription termination by RNA polymerase III. *Genetics* **125,** 293–303 (1990).
141. Shaaban, S. A., Krupp, B. M. & Hall, B. D. Termination-altering mutations in the second-largest subunit of yeast RNA polymerase III. *Molecular and Cellular Biology* **15,** 1467–1478 (1995).
142. Shaaban, S. A., Bobkova, E. V., Chudzik, D. M. & Hall, B. D. In vitro analysis of elongation and termination by mutant RNA polymerases with altered termination behavior. *Molecular and Cellular Biology* **16,** 6468–6476 (1996).
143. Rijal, K. & Maraia, R. J. RNA polymerase III mutants in TFIIFα-like C37 that cause terminator readthrough with no decrease in transcription output. *Nucleic Acids Research* **41,** 139–155 (2013).
144. Rijal, K. & Maraia, R. J. Active Center Control of Termination by RNA Polymerase III and tRNA Gene Transcription Levels In Vivo. *PLoS Genetics* **12,** e1006253 (2016).
145. Fernández-Tornero, C. *et al.* Insights into Transcription Initiation and Termination from the Electron Microscopy Structure of Yeast RNA Polymerase III. *Molecular Cell* **25,** 813–823 (2007).
146. Kireeva, M. L., Komissarova, N., Waugh, D. S. & Kashlev, M. The 8-Nucleotide-long RNA:DNA Hybrid Is a Primary Stability Determinant of the RNA Polymerase II Elongation Complex. *Journal of Biological Chemistry* **275,** 6530–6536 (2000).
147. Martin, F. H. & Tinoco, I. DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Research* **8,** 2295–2300 (1980).
148. Maraia, R. J. & Rijal, K. A transcriptional specialist resolved. *Nature* **528,** 204–205 (2015).
149. Turowski, T. W. *et al.* Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome Research* **26,** 933–944 (2016).
150. Rivosecchi, J. *et al.* Senataxin homologue Sen1 is required for efficient termination of RNA polymerase III transcription. *The EMBO Journal* **38,** e101955 (2019).
151. Kühlbrandt, W. The Resolution Revolution. *Science* **343,** 1443–1444 (2014).
152. Fernandez-Leiro, R. & Scheres, S. H. W. Unravelling biological macromolecules with cryo-electron microscopy. *Nature* **537,** 339–346 (2016).
153. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P. & Ranson, N. A. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* **100,** 3–15 (2016).

154. Dubochet, J. & McDowall, A. Vitrification Of Pure Water For Electron Microscopy. *Journal of Microscopy* **124,** 3–4 (1981).
155. Adrian, M., Dubochet, J., Lepault, J. & McDowall, A. W. Cryo-electron microscopy of viruses. *Nature* **308,** 32–36 (1984).
156. Frank, J., Verschoor, A. & Boublik, M. Computer averaging of electron micrographs of 40S ribosomal subunits. *Science* **214,** 1353–1355 (1981).
157. Frank, J. Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nature Protocols* **12,** 209–212 (2017).
158. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **180,** 519–530 (2012).
159. Scheres, S. H. W. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods in Enzymology* **579,** 125–157 (2016).
160. McMullan, G., Chen, S., Henderson, R. & Faruqi, A. R. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* **109,** 1126–1143 (2009).
161. Brilot, A. F. *et al.* Beam-induced motion of vitrified specimen on holey carbon film. *Journal of Structural Biology* **177,** 630–637 (2012).
162. Girbig, M. *et al.* Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. *Nature Structural and Molecular Biology* **28,** 210–219 (2021).
163. Chen, J., Noble, A. J., Kang, J. Y. & Darst, S. A. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *Journal of Structural Biology: X* **1,** 100005 (2019).
164. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14,** 290–296 (2017).
165. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nature Methods* **16,** 1146–1152 (2019).
166. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7,** e42166 (2018).
167. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7,** e36861 (2018).
168. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology* **213,** 107702 (2021).
169. Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Computational Biology* **7,** e1002195 (2011).
170. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Research* **47,** D427–D432 (2019).
171. Kim, M. K. *et al.* Assembly of SNAPc, Bdp1, and TBP on the U6 snRNA Gene Promoter in *Drosophila melanogaster*. *Molecular and Cellular Biology* **40,** e00641–19 (2020).
172. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43,** D512–D520 (2015).

173. Hwang, S., Gou, Z. & Kuznetsov, I. B. DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins. *Bioinformatics* **23,** 634–636 (2007).
174. Palian, B. M. *et al.* Maf1 Is a Novel Target of PTEN and PI3K Signaling That Negatively Regulates Oncogenesis and Lipid Metabolism. *PLoS Genetics* **10** (ed Kim, S. K.) e1004789 (2014).
175. Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *Nature Methods* **17,** 923–927 (2020).
176. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6,** 343–345 (2009).
177. Weissmann, F. *et al.* biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proceedings of the National Academy of Sciences* **113,** E2564–E2569 (2016).
178. Leśniewska, E. & Boguta, M. Novel layers of RNA polymerase III control affecting tRNA gene transcription in eukaryotes. *Open Biology* **7,** 170001 (2017).
179. Wang, Q. *et al.* Structural insights into transcriptional regulation of human RNA polymerase III. *Nature Structural & Molecular Biology* **28,** 220–227 (2021).
180. Ramsay, E. P. *et al.* Structure of human RNA polymerase III. *Nature Communications* **11,** 6409 (2020).
181. Li, L. *et al.* Structure of human RNA polymerase III elongation complex. *Cell Research* **31,** 791–800 (2021).
182. Wu, L. *et al.* Novel Small-Molecule Inhibitors of RNA Polymerase III. *Eukaryotic Cell* **2,** 256–264 (2003).
183. Wang, X., Rusin, A., Walkey, C. J., Lin, J. J. & Johnson, D. L. The RNA polymerase III repressor MAF1 is regulated by ubiquitin-dependent proteasome degradation and modulates cancer drug resistance and apoptosis. *Journal of Biological Chemistry* **294,** 19255–19268 (2019).
184. Braglia, P., Percudani, R. & Dieci, G. Sequence Context Effects on Oligo(dT) Termination Signal Recognition by *Saccharomyces cerevisiae* RNA Polymerase III. *Journal of Biological Chemistry* **280,** 19551–19562 (2005).
185. Valdar, W. S. Scoring residue conservation. *Proteins: Structure, Function, and Genetics* **48,** 227–241 (2002).
186. Granneman, S., Kudla, G., Petfalski, E. & Tollervey, D. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proceedings of the National Academy of Sciences* **106,** 9613–9618 (2009).
187. Richard, P. & Manley, J. L. Transcription termination by nuclear RNA polymerases. *Genes & Development* **23,** 1247–1269 (2009).
188. Hamada, M., Sakulich, A. L., Koduru, S. B. & Maraia, R. J. Transcription Termination by RNA Polymerase III in Fission Yeast. *Journal of Biological Chemistry* **275,** 29076–29081 (2000).
189. Goodarzi, H. *et al.* Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165,** 1416–1427 (2016).

190. Keeling, P. J. & Burki, F. Progress towards the Tree of Eukaryotes. *Current Biology* **29,** R808–R817 (2019).
191. Steinmetz, E. J., Conrad, N. K., Brow, D. A. & Corden, J. L. RNA-binding protein Nrd1 directs poly(A)-independent 3-end formation of RNA polymerase II transcripts. *Nature* **413,** 327–331 (2001).
192. Porrua, O. & Libri, D. A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nature Structural & Molecular Biology* **20,** 884–891 (2013).
193. Moreno-Morcillo, M. *et al.* Solving the RNA polymerase I structural puzzle. *Acta Crystallographica Section D Biological Crystallography* **70,** 2570–2582 (2014).
194. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* **152,** 36–51 (2005).
195. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7,** 253–267 (2020).
196. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *Journal of Structural Biology* **157,** 281–287 (2007).
197. Rosenthal, P. B. & Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology* **333,** 721–745 (2003).
198. Vilas, J. L. *et al.* MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps. *Structure* **26,** 337–344.e4 (2018).
199. De la Rosa-Trevín, J. *et al.* Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *Journal of Structural Biology* **195,** 93–99 (2016).
200. Morin, A. *et al.* Collaboration gets the most out of software. *eLife* **2,** e01456 (2013).
201. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods* **17,** 1214–1221 (2020).
202. Ramírez-Aportela, E. *et al.* Automatic local resolution-based sharpening of cryo-EM maps. *Bioinformatics* **36,** 765–772 (2019).
203. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D Structural Biology* **75,** 861–877 (2019).
204. Bienert, S. *et al.* The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Research* **45,** D313–D319 (2017).
205. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10,** 845–858 (2015).
206. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography* **66,** 486–501 (2010).
207. Casañal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Science* **29,** 1055–1064 (2020).

208. Nicholls, R. A., Long, F. & Murshudov, G. N. Low-resolution refinement tools in REFMAC 5. *Acta Crystallographica Section D Biological Crystallography* **68,** 404–417 (2012).
209. Croll, T. I. ISOLDE : a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D Structural Biology* **74,** 519–530 (2018).
210. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* **27,** 14–25 (2018).

Appendix

Table A1. Human Pol III protein subunits, gene IDs, and sources of ordered cDNA clones

Protein	Uniprot ID	Gene	NCBI ID	Source	Clone ID
RPC1	O14802	POLR3A	11128	BioCat	BC041089-TCH1003-GVO-TRI
RPC2	Q9NW08	POLR3B	55703	BioCat	BC046238-TCH1003-GVO-TRI
RPC3	Q9BUI4	POLR3C	10623	DNASU	HsCD00674661
RPC4	P05423	POLR3D	661	DNASU	HsCD00287815
RPC5	Q9NVU0	POLR3E	55718	DNASU	HsCD00514436
RPC6	Q9H1D9	POLR3F	10621	DNASU	HsCD00041908
RPC7	O15318	POLR3G	10622	DNASU	HsCD00080109
RPC8	Q9Y535	POLR3H	171568	DNASU	HsCD00073948
RPC9	O75575	CRCP	27297	BioCat	BC105808-TCH1003-GVO-TRI
RPC10	Q9Y2Y1	POLR3K	51728	DNASU	HsCD00042020
RPAC1	O15160	POLR1C	9533	DNASU	HsCD00288772
RPAC2	P0DPB6	POLR1D	51082	DNASU	HsCD00444528
RPABC1	P19388	POLR2E	5434	DNASU	HsCD00288908
RPABC2	P61218	POLR2F	5435	DNASU	HsCD00073732
RPABC3	P52434	POLR2H	5437	DNASU	HsCD00079996
RPABC4	P53803	POLR2K	5440	DNASU	HsCD00040026
RPABC5	P62875	POLR2L	5441	DNASU	HsCD00004472

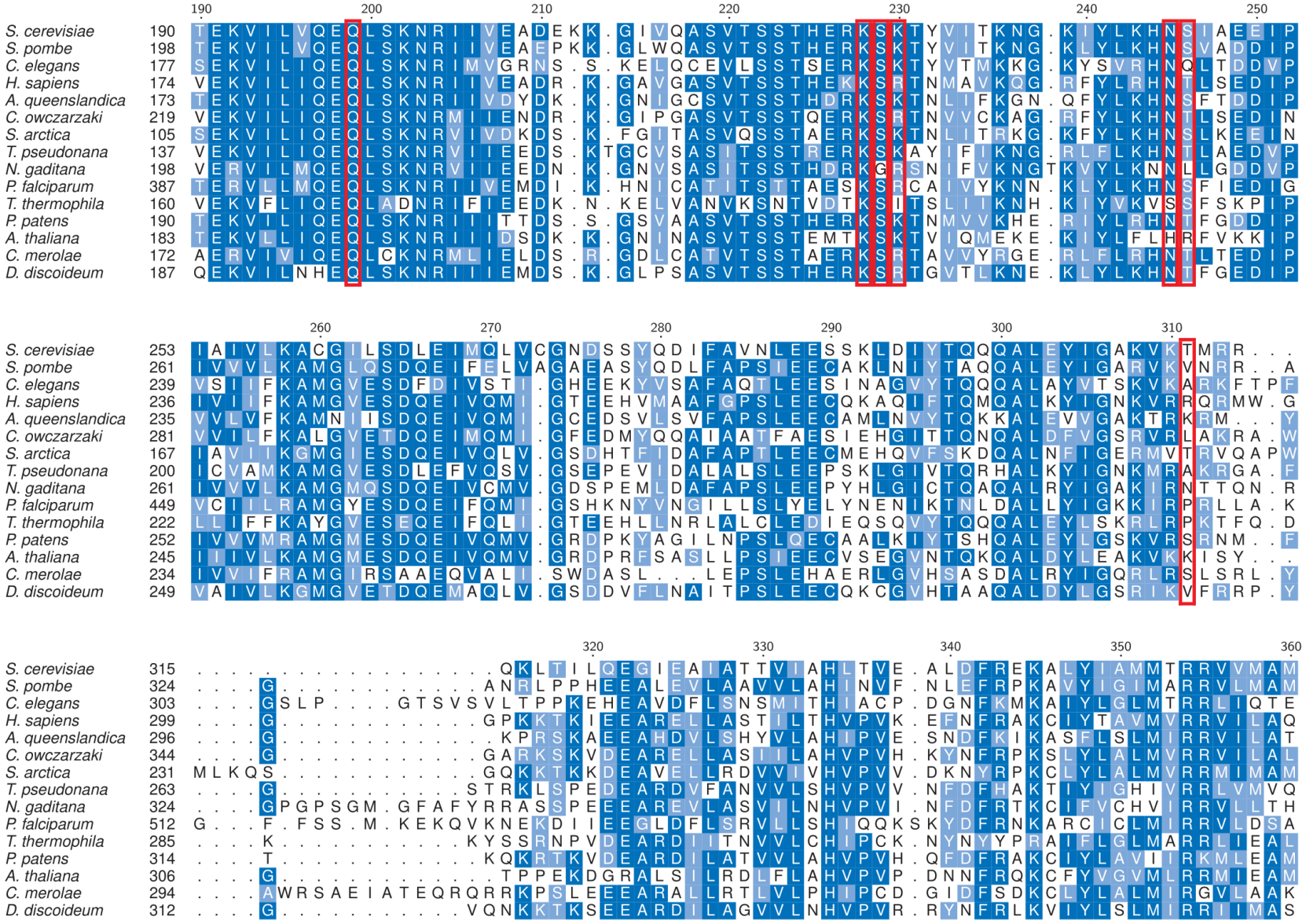

Fig. A1. Multiple sequence alignment of subunit RPC2 (C128). The sequence range is 190 to 490 in *S. cerevisiae*. Residues that form H-bonds with the NT-strand are highlighted in red. Covered are five eukaryotic supergroups: ophistokonts (*S. cerevisiae, S. pombe, C. elegans, H. sapiens, A. queenslandica, C. owczarzaki, S. arctica*); stramenopiles (*T. pseudonana, N. gaditana*); alveolata (*P. falciparum, T. thermophila*); archaeplastids (*P. patens, A. thaliana, C. merolae*); amoebozoans (*D. discoideum*).

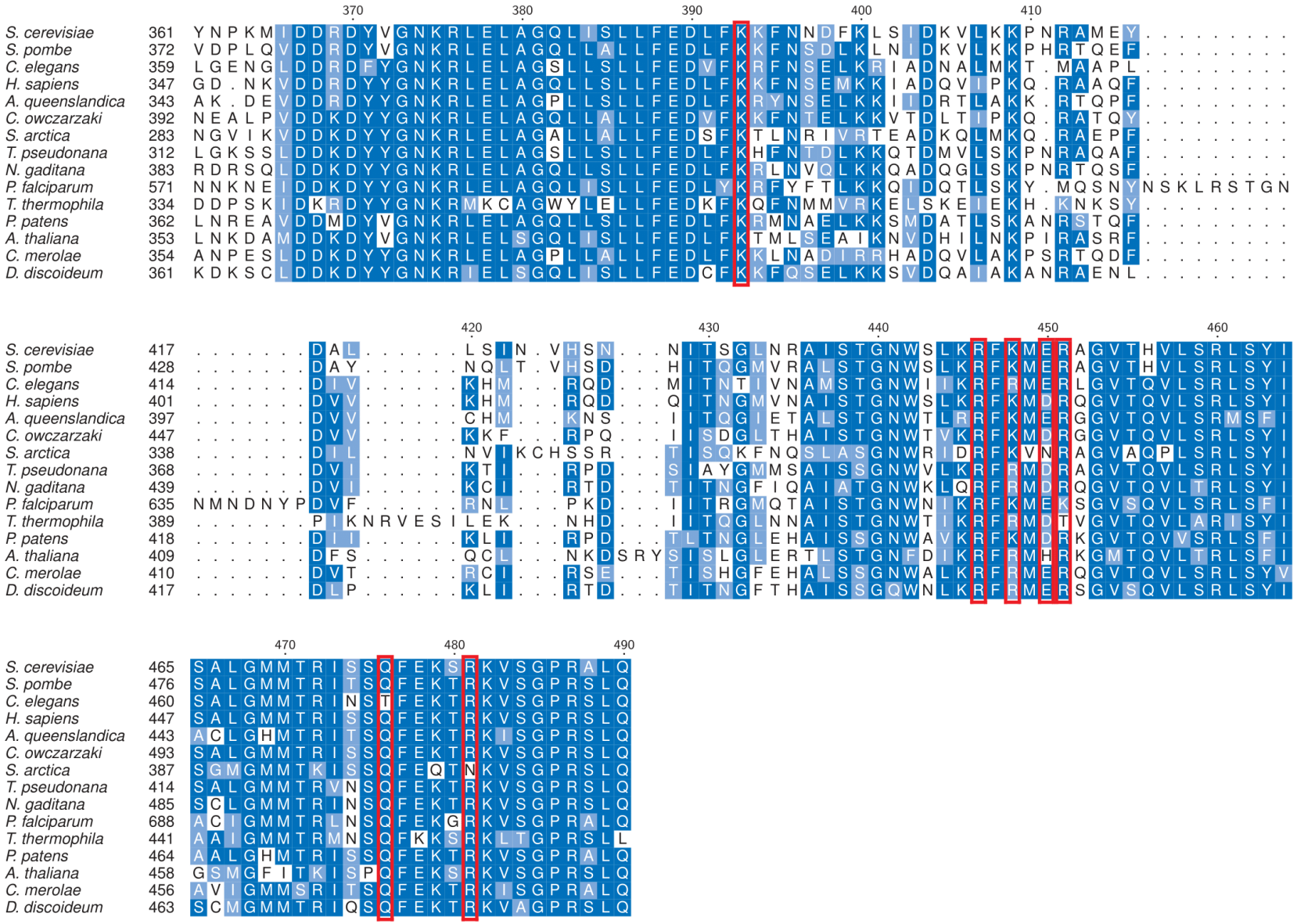

Fig. A1. continued

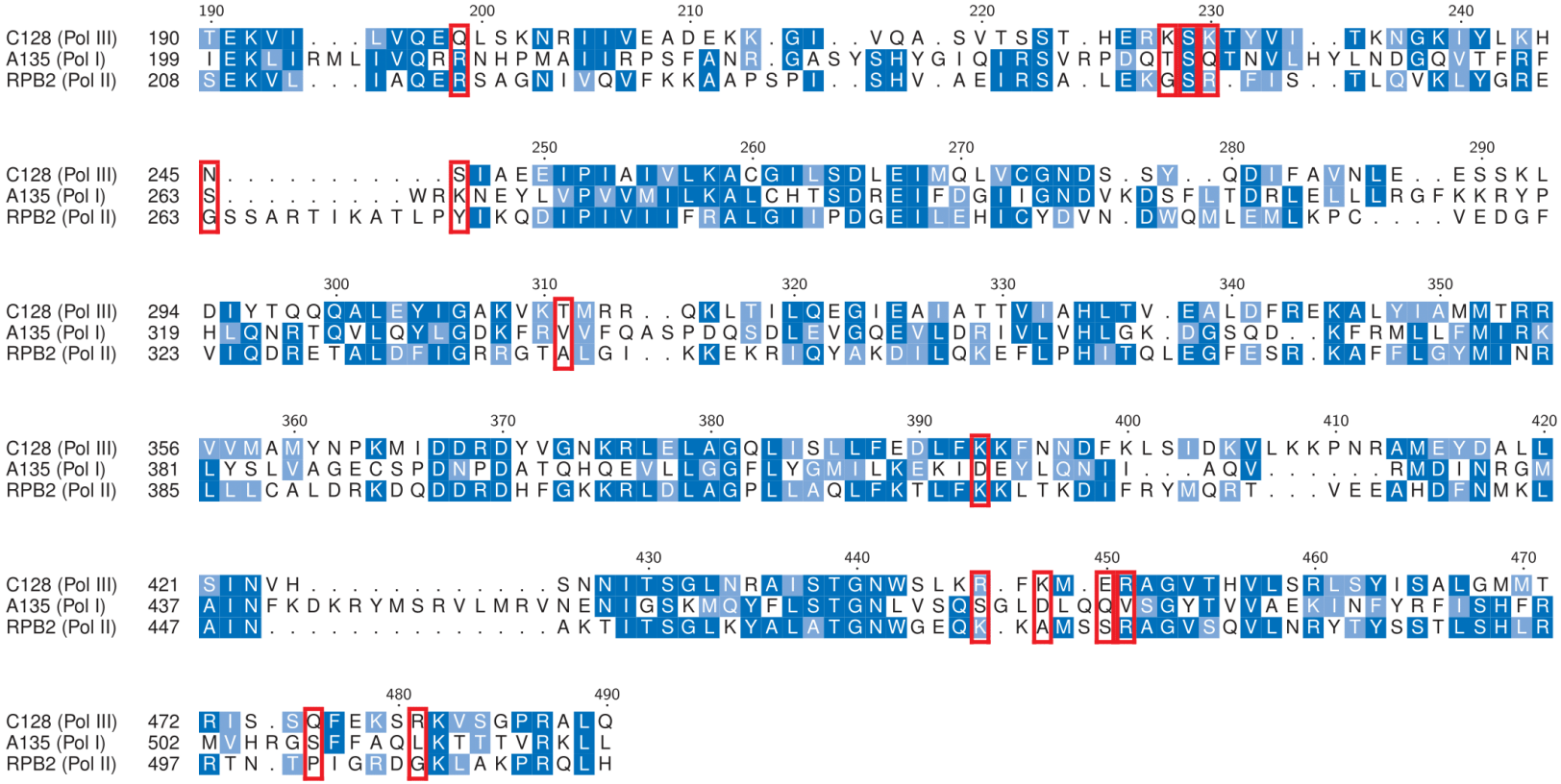

Fig. A2. Multiple sequence alignment of *S. cerevisiae* C128 (Pol III) and it paralogs A135 (Pol I) and RBP1 Pol II. The sequence range is 190 to 490 in C128. Residues that form H-bonds with the NT-strand in the Pol III PTC are highlighted in red.