Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Biologie und vorklinische Medizin
der Universität Regensburg

Structural and functional characterization of the human RNA polymerase I transcription system

vorgelegt von

Julia Daiß

aus

Karlsruhe

im Jahr 2023

Das Promotionsgesuch wurde eingereicht am:
27.9.2023

Die Arbeit wurde angeleitet von:
Prof. Dr. Christoph Engel

Unterschrift

Summary

Transcription of DNA is a fundamental mechanism of every living cell and is carried out by at least three nuclear DNA-dependent RNA polymerases (Pols) in eukaryotes. Among these, Pol I is dedicated to the transcription of one specific gene. The resulting 47S ribosomal RNA (rRNA) precursor is further processed to 18S, 5.8S and 28S rRNA and incorporated into the cellular machinery for protein translation, the ribosome. Pol I is a major determinant of cellular growth. Mutations or deregulation of the Pol I transcription system have been found to be linked to severe (neuro-)developmental diseases and many types of cancer, pointing to the direct link between human health and proper Pol I transcription. The last years, the spotlight of modern medicine turned on the Pol I system, investigating different drugs targeting Pol I transcription as cancer-treatment strategy. Nevertheless, basic Pol I research has been mainly performed in the yeast system and a detailed understanding on a mechanistic, structural, and functional level was still missing for the human enzyme. My overall research aim was to investigate which mechanisms known from the yeast Pol I system can be transferred to human and which of the features are species-specific. In higher eukaryotes, regulatory variants dependent on tissue type, developmental state and cell-cycle stage add additional layers of complexity. Hence, it remained unclear how Pol I structurally and functionally adapted to the increased regulatory demands in human cells.

To tackle these questions, an *in vitro* system for human Pol I was needed. A human cell line was established, carrying an endogenous GFP-tag on the largest subunit RPA1 of Pol I. The modified cell line allowed to track Pol I over the cell cycle and to purify the human enzyme to high purity and quality enabling a structure-function analysis. The predominant configuration of Pol I was shown to be a 13-subunit enzyme including a single-subunit stalk, as present in humans, but not in yeast. *In vitro* functional assays using a minimal DNA-RNA scaffold showed cleavage and elongation activity for hPol I. Comparing hPol I to its yeast counterpart revealed a reduced proof-reading activity on the human enzyme due to a reduced cleavage and proof-reading activity during nucleotide addition. Using cryogenic Electron Microscopy (cryo-EM), a single-particle reconstruction of monomeric hPol I was obtained at 4.09 Å overall resolution allowing the generation of a molecular model. The architecture of the enzyme is conserved between yeast and human. The reconstruction of monomeric hPol I showed connected density for most of the core polymerase and the RPA49/34 dimerization domain. However, the clamp subdomain, the stalk, the C-ribbon of subunit RPA12, the linker and tWH of subunit RPA49, and the human-specific C-terminal region of subunit RPA34 stayed flexible. Enhanced clamp and stalk flexibility was a divergent feature between yeast and human and might suggest that there is no defined state for this region in monomeric hPol I. The foot domain of subunit RPA1 diverges between yeast and human, comprising an additional domain termed ‘dock II’ in humans. Dock II resembles a truncated HMG-box that we found to be incapable of DNA-binding. *In situ* modelling, *in vitro* pulldowns and similarity indicated that dock II may serve as a downstream-transcription factor binding platform. This included the possibility of Top2a recruitment, since re-evaluation of Chromatin Immuno-Precipitation (ChIP) data revealed that Top2a was enriched at the rDNA locus. Interestingly, peaks of high Top2a occupancy overlapped with peaks of the metazoan-specific, six HMG-box-containing transcription factor UBF. Indeed, there was Top2a-UBF interaction *in vitro*, indicating functional cooperativity.

This thesis not only provides evidence for overall functional and structural conservation throughout evolution, but additionally uncovers organism-specific features.

Zusammenfassung

Die Transkription von DNA zu RNA ist einer der essentiellen Mechanismen jeder lebenden Zelle und wird in Eukaryoten von mindestens drei verschiedenen nuklearen DNA-abhängigen RNA Polymerasen (Pol) katalysiert. Pol I ist dabei auf einen Promoter spezialisiert und transkribiert nur ein Gen, welches für die 47S Vorläufer-rRNA (ribosomale RNA) codiert. Die daraus generierte RNA wird weiter zu 18S, 5.8S und 28S rRNA prozessiert. Gemeinsam mit 5S rRNA und den ribosomalen Proteinen wird daraus der zelluläre Komplex zur Proteintranslation, das Ribosom, gebildet. Aufgrund ihrer wichtigen Funktion, der Synthese von rRNAs, ist Pol I ein entscheidender Regulator des zellulären Wachstums. Diese Relevanz führt dazu, dass Mutationen und Fehlregulationen dieses Transkriptionssystems zu schweren (neuro-)degenerativen Erkrankungen führen können und mit verschiedenen Krebsarten in Verbindung stehen. Durch die direkte Korrelation zwischen korrekter Pol I Transkription und menschlicher Gesundheit ist das Pol I System in den letzten Jahren in den Fokus der modernen Medizin gerückt. Aktuell werden erste Medikamente, welche die Pol I Transkription regulieren, entwickelt und getestet. Der Großteil der bisherigen Pol I Forschung wurde im Hefe-System durchgeführt, weshalb detaillierte mechanistische, funktionelle und strukturelle Analysen des menschlichen Enzyms noch gefehlt haben. Das Hauptziel dieser Arbeit war zu untersuchen, welche Mechanismen und Eigenschaften der Hefe Pol I auf das menschliche Enzym übertragen werden können und welche spezifisch für die jeweilige Art sind. Außerdem vervielfacht sich in höheren Eukaryoten, wie dem Menschen, die Komplexität, da das Pol I System zusätzlich je nach Gewebetyp, Entwicklungsstatus und Zellzyklus-Phase reguliert werden muss. Deshalb wurden die strukturellen und funktionellen Eigenschaften des menschlichen Enzyms näher untersucht, um herauszufinden, welche Anpassungen an die vielschichtigen Anforderungen stattgefunden haben.

Um diese Fragen beantworten zu können, musste zunächst ein System etabliert werden, welches die *in vitro* Analysen und die Aufreinigung von humaner Pol I (hPol I) erlaubt. Dafür wurde zunächst eine humane Zelllinie so verändert, dass die größte Untereinheit der Polymerase RPA1 einen GFP (*green flourescent protein*)-Tag trägt. Dadurch konnte zum einen das humane Enzym in Zellen visualisiert und im Verlauf des Zellzyklus beobachtet werden, zum anderen konnte hPol I mit hoher Qualität und Reinheitsgrad aufgereinigt werden. Es konnte gezeigt werden, dass die vorherrschende Zusammenstellung von Pol I aus 13 Untereinheiten besteht und der *Stalk* nur von einer einzelnen Untereinheit gebildet wird, was auch auf das menschliche Enzym zutrifft, allerdings nicht im Hefe-Modellorganismus. Funktionsanalysen zeigten hPol I Aktivität, sowohl bei Spaltung (*cleavage*) und Elongation. Vergleicht man diese Aktivität mit derjenigen von Pol I aus Hefe, konnte eine verringerte Korrekturleseaktivität (*proofreading*) des menschlichen Enzyms erkannt werden, welche sich auf die reduzierte Spaltungsaktivität und/oder auf einen erhöhten tolerierten Einbau falscher Nukleotide zurückführen lässt. Mithilfe der Kryoelektronenmikroskopie (Kryo-EM) konnte die Dichte von hPol I im monomeren Zustand mit einer Auflösung von 4.09 Å bestimmt und das Modell erstellt werden. Damit konnte gezeigt werden, dass die allgemeine Architektur zwischen Hefe und Mensch konserviert ist. Die Rekonstruktion der menschlichen Pol I besteht aus zusammenhängender Dichte für die Kern-Polymerase (*core polymerase*) sowie der RPA49/34 Dimerisierungsdomäne. Dennoch waren einige Domänen flexibel und daher unsichtbar: die *Clamp*-Domäne, der *Stalk*, der *C-ribbon* der Untereinheit RPA12, der Linker und die *tWH* von RPA49, sowie die C-terminale Region von RPA34. Die Flexibilität der menschlichen *Clamp* und *Stalk* Regionen

war im Unterschied zur Hefe Pol I deutlich stärker ausgeprägt, was auf eine fehlende definierte räumliche Lage dieser Regionen im monomeren Zustand hinweisen könnte. Ein weiterer Unterschied der humanen Pol I im Gegensatz zur Hefe ist eine zusätzliche Domäne im *Foot* der Untereinheit RPA1, welche als *Dock II* bezeichnet wurde. *Dock II* ähnelt einer HMG-box und ist nicht DNA-bindend. *In situ* Modellierungen, *in vitro Pulldowns* and Ähnlichkeiten weisen darauf hin, dass diese *Metazoa*-spezifische Domäne möglicherweise eine zusätzliche Plattform für die Bindung von Transkriptionsfaktoren sein könnte, wie beispielsweise der Topoisomerase 2a (Top2a). Weitergehende Analysen bereits veröffentlichter Chromatin Immuno-Präzipitation (ChIP) Daten zeigten, dass Top2a tatsächlich am rDNA-Lokus angereichert ist. Interessanterweise konnte dabei auch festgestellt werden, dass einige Stellen erhöhter lokaler Top2a Konzentration mit erhöhter Binding des *Metazoa*-spezifischen und sechs HMG-Boxen umfassenden Transkriptionsfaktor UBF überlappen. *In vitro* Untersuchungen konnten zeigen, dass eine direkte Interaktion der beiden Proteine Top2a und UBF möglich ist, was auf eine funktionelle Kooperativität hindeutet.

Diese Arbeit bietet somit nicht nur Einblicke in die funktionelle und strukturelle Konservierung im Verlauf der Evolution, sondern auch in die spezifischen Eigenschaften der verschiedenen Organismen.

Publications

Parts of the human Pol I analysis were published in:

Daiß, J. L., Pilsl, M., Straub, K., Bleckmann, A., Höcherl, M., Heiss, F. B., Abascal-Palacios, G., Ramsay, E. P., Tlučková, K., Mars, J.-C., Fürtges, T., Bruckmann, A., Rudack, T., Bernecky, C., Lamour, V., Panov, K., Vannini, V., Moss, T. & Engel, C.✉ The human RNA polymerase I structure reveals an HMG-like docking domain specific to metazoans. *Life science alliance* **5**; 10.26508/lsa.202201568 (2022).

Contributions:
J.L.D. designed and performed genome-editing to create the stable HeLa-POLR1A cell line. C.E., M.P., J.L.D., E.P.R. and G.A.-P. performed initial small-scale purifications of human Pol I. J.L.D. optimized suspension cell culture of HeLa-POLR1A, performed large-scale purifications of human Pol I, and carried out transcription assays. A.B. und J.L.D. performed confocal microscopy. J.L.D. prepared negatively stained EM sample and C.E., M.P., F.B.H., and J.L.D. performed negative stain EM data collection and processing. J.L.D., M.P., E.P.R., G.A.-P., K.T., C.B. and F.H. carried out cryo-EM sample preparation and optimization as well as cryo-EM screening. J.L.D., M.P., E.P.R., G.A.-P., C.E., and A.V. performed cryo-EM data collection and processing. C.E. carried out model building. G.A.-P., E.P.R., A.V., and J.L.D performed mapping of disease-causing mutations. K.S. carried out phylogenetic analysis. J.L.D. performed human RRN3 purification, M.P. and J.L.D. *S. cerevisiae* Pol I purification, J.L.D. and M.H. purification of recombinant heterodimer and dock II versions, M.P. yeast Rrn3 and F.B.H. *S. pombe* Pol I purifications. V.L. supplied full-length and ΔCTD Top2a and K.P. purified recombinant fUBF. A.B. carried out mass spectrometry analysis. J.L.D. performed DNA-binding analysis of heterodimer and dock II versions via EMSAs. J.-C.M. and T.M. executed re-analysis of published ChIP-Seq data. J.L.D. analyzed dock II-Top2a interactions and K.P. performed Top2a co-IPs and UBF-Top2a pulldowns. T.F. and T.R. carried out molecular modelling of Top2-Pol I/dock II/UBF interactions. C.E. designed and supervised research and wrote the manuscript draft. C.E. and J.L.D. edited and reviewed the manuscript with input from all authors.

Parts of the human Pol III analysis were published in:

Ramsay, E. P.*, Abascal-Palacios, G.*, **Daiß, J. L.***, King, H., Gouge, J., Pilsl, M., Beuron, F., Morris, E., Gunkel, P., Engel, C.✉ & Vannini, A.✉ Structure of human RNA polymerase III. *Nature communications* **11**, 6409; 10.1038/s41467-020-20262-5 (2020).

Contributions:
C.E. and P.G. designed and performed genome-editing experiments. P.G. created the stable POLR1C-sfGFP HeLa cell line. C.E., M.P. and J.L.D. performed initial small-scale purifications of human Pol III. J.L.D. optimized suspension cultivation of RPAC1-GFP cells, carried out transcription assays and confocal microscopy. G.A.-P. and E.P.R. performed large-scale purification of human Pol III and prepared samples for EM analysis and collected cryo-EM data. F.B. carried out cryo-EM sample preparation, screening and sample collection, and E.M. helped with preliminary evaluation of data. E.P.R. performed cryo-EM data processing. G.A.-P. performed X-ray crystallography experiments and

analyzed the data. J.G. collected SAXS data. E.P.R. and G.A.-P. analyzed SAXS data. H.K. performed experiments in living cells. G.A.-P., E.P.R., H.K. and A.V. interpreted the data. A.V. wrote the manuscript with input from all the authors. C.E. and A.V. designed and supervised research.

Parts of the hPol I introduction and discussion were published in:

Daiß, J. L., Griesenbeck, J.✉, Tschochner, H.✉ & Engel, C.✉ Synthesis of the ribosomal RNA precursor in human cells: mechanisms, factors and regulation. *Biological chemistry*; 10.1515/hsz-2023-0214 (2023).

Parts of the hPol III introduction and discussion were published in:

Wang, Q.*, **Daiß, J. L.***, Xu, Y.✉ & Engel, C.✉ Snapshots of RNA polymerase III in action - A mini review. *Gene* **821,** 146282; 10.1016/j.gene.2022.146282 (2022).

Additional work:

Heiss, F. B., **Daiß, J. L.**, Becker, P. & Engel, C.✉ Conserved strategies of RNA polymerase I hibernation and activation. *Nature communications* **12,** 758; 10.1038/s41467-021-21031-8 (2021).

Contributions:
F.H. carried out Sp Pol I purification and characterization, prepared cryo-EM grids and carried out sequence analysis. F.H. and C.E. processed cryo-EM data. J.D. carried out functional elongation/cleavage assays. P.B. and C.E. built Sp Pol I models. C.E. designed and supervised research, established strains and purification protocols and prepared the manuscript with input from all authors.

Rausch, C., Zhang, P.✉, Casas-Delucchi, C. S.*, **Daiß, J. L.***, Engel, C., Coster, G., Hastert, F. D., Weber, P. & Cardoso, M. C.✉ Cytosine base modifications regulate DNA duplex stability and metabolism. *Nucleic acids research* **49**, 12870-12894; 10.1093/nar/gkab509 (2021).

Contribution:
C.R., P.Z. and M.C.C. conceived the project. C.R. generated final figures and C.R. and M.C.C. wrote the manuscript. All authors revised, commented and agreed on the manuscript. C.R. performed and analyzed all cell-based assays and contributed to *in vitro* PCR, HRM and run-off transcription experiments. P.Z. performed and analyzed *in vitro* PCR and HRM experiments. F.H. performed and analyzed *in vitro* run-off transcription experiments. P.W. contributed to the ES cell cycle distribution, S-phase progression experiments and performed molecular combing experiments. C.C.-D. and G.C. performed and analyzed *in vitro* yeast replication assays. J.L.D. and C.E. performed and analyzed *in vitro* yeast transcription assays.

* contributed equally
✉ corresponding authors

Table of contents

1 Introduction

1.1 Importance of RNA polymerase I

The central dogma of molecular biology includes the flow of genetic information from DNA over RNA to complete protein[1]. Essential to this process, transcription from DNA to RNA is enabled by DNA-dependent RNA polymerases (Pols) and translation from RNA to proteins is performed by ribosomes with the help of aminoacylated tRNAs (transfer RNAs). In eukaryotes, at least three specific RNA polymerases have evolved (in plants five), which are highly conserved throughout evolution and are located within the nucleus[2–4].

Among the three Pols, overall structure and function is conserved as well as the steps of the transcription cycle. Nevertheless, transcription associated factors vary, thereby contributing to enzyme specificity and specializations[5–7]. Pol II decodes protein-coding genes to mRNA (messenger RNA) and produces many small regulatory RNAs, such as miRNAs (micro RNAs)[8]. Pol III produces tRNAs, needed during translation from mRNA to protein, 5S rRNA (ribosomal RNA) and many long non-coding RNAs such as U6 RNA or 7SK RNA. Pol III is the largest of the three Pols, composed of 17 subunits[9]. Unlike the previously described RNA polymerases, Pol I transcribes only one single gene, the rRNA precursor (pre-rRNA) gene[10–12]. In mammals, the resulting 47S pre-rRNA is processed to result in the three rRNAs 28S, 18S, and 5.8S[10]. Those three rRNAs transcribed by Pol I form together with the 5S rRNA produced by Pol III and many ribosomal proteins (r-proteins) the ribosome, the cellular machinery responsible for translation of proteins[13]. The amount of ribosomes is a critical determinant of cell growth[10] and rRNA transcription by Pol I is a key regulatory step during ribosome biogenesis[14–16]. Thus, Pol I transcription must be tightly regulated and tuned to the growth status of the cell as well as to nutrient availability[10,17]. In humans, Pol I transcription accounts for about 35 % of RNA synthesis and rRNAs account for about 80 % of total cellular RNA[11,18], which are the main component of up to 10 million ribosomes within one cell[19]. Furthermore, considering the close link between Pol I transcription and cellular growth, it's not surprising that many proto-oncogenes and tumor suppressors are involved in the regulation of this process[11,20,21]. Deregulation is observed in many different types of cancer and a range of severe developmental and neurological disorders[18,22].

1.2 rDNA organization

Pol I transcription, pre-rRNA processing and the first steps of ribosome biogenesis take all place within a distinct sub-compartment of the nucleus, the nucleolus[10]. The mammalian nucleus can harbor more than one nucleolus which originate from different nucleolar organizer regions (NORs)[10,11,23,24]. In humans, five NORs exist in the haploid genome, located at the acrocentric chromosomes 13, 14, 15, 21, and 22, all together containing 200 - 300 ribosomal genes[11,25] (Figure 1A). Each NOR consists of many ribosomal DNA (rDNA) gene repeats, mainly in head-to-tail orientation (Figure 1B), and has the potential to form a nucleolus, or stays heterochromatic and transcriptionally inactive[10,11,19]. In non-transformed human cell lines, it was observed that all NORs are transcriptionally active, whereas in

cancer cell lines there is evidence that individual NORs might be transcriptionally silent[26]. Ultrastructural analysis of active nucleoli revealed nucleolar sub-compartments. Fibrillar centers and associated dense fibrillar centers are the location of pre-rRNA synthesis and co-transcriptional rRNA processing. Further steps of large and small ribosomal subunit maturation occur in outer granular centres[11,27,28]. The 5S rRNA gene array, transcribed by Pol III, is located on chromosome 1 and not physically connected to the NORs in human cells, but may associate with the periphery of nucleoli and hence facilitate 5S rRNA import[11,29–31]. 47S pre-rRNA processing includes co- and posttranscriptional endo- and exonucleolytic cleavage and assembly with ribosome biogenesis factors and r-proteins. The majority of the multiple steps leading to mature ribosomes takes place in the nucleolus, followed by further steps in the nucleus and in the cytoplasm [13,19]. Mature ribosomes consist of the 60S large subunit comprising 28S, 5.8S, and 5S rRNA, as well as around 50 r-proteins and the 40S small ribosomal subunit including 18S rRNA as well as about 33 r-proteins, which together assemble to the 4 MDa cellular protein-producing machinery[11,19].

Pol I transcription from active rDNA genes is highly processive, leading to a dense packing of Pol I enzymes in the transcribed rRNA gene region[32]. Visualization of Pol I transcription and co-transcriptional processing was achieved for many organisms by electron microscopy of so-called 'Miller chromatin spreads'[33,34] (Figure 1C). In Miller spreads, also referred to as 'Christmas trees', the rDNA bound by many elongating Pol I enzymes forms the stem of the tree, while the synthesized pre-rRNA extends from the stem forming the branches, the terminal ball-like structures represent pre-rRNA assembly intermediates[13,35,36]. The mammalian rDNA gene consists of a spacer promoter with a termination site followed by enhancer repeats, another termination site, the actual promoter with the transcription start site (TSS), the transcribed 47S pre-rRNA followed by several termination sites at the 3' end of the gene, and an intergenic spacer (IGS) spanning the region between the genes[37]. The transcribed region is composed of the 5' external transcribed spacer (ETS), 18S rRNA, internal transcribed spacer (ITS) 1, 5.8S rRNA, ITS2, 28S rRNA and 3' ETS[13,37] (Figure 1B).

Pol I recognizes only one specific promoter which is composed of a core promoter located around the TSS (around -45 to +20) and the upstream control element (UCE) also referred to as upstream promoter element (UPE) (around -234 to -107)[11,38]. The DNA sequence of the promoter region is poorly conserved among species which may explain the species specificity of transcription initiation factors of the Pol I transcription system[11]. Instead, DNA-structure rather than sequence has been suggested to be the common feature of Pol I promoters including a conserved DNA bendability and meltability which might assist transcription initiation[39].

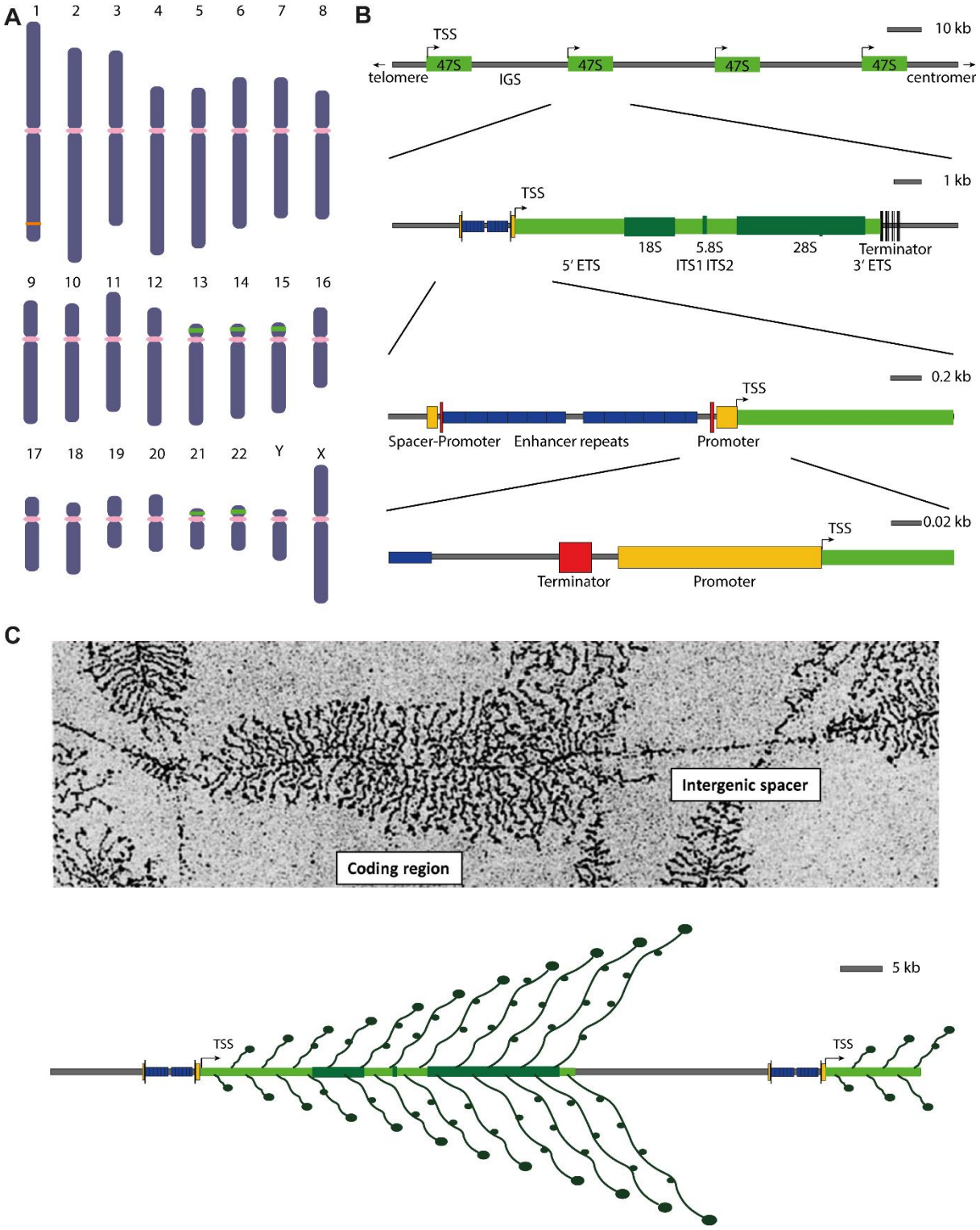


Figure 1 The genomic rDNA locus in human cells

A Schematic representation of human chromosomes with centromeres (labeled in pink). The genomic loci of 47S rDNA repeats (labeled in green) and of 5S rDNA repeat (labeled in orange) are pointed out. **B** Zoom-in into the genomic rDNA locus: on top visualization of the 47S pre-rRNA gene with the transcription start site (TSS; labeled with an arrow) and the intergenic spacer region (IGS, grey) in relation to telomere and centromere. Underneath: Schematic representation of the rDNA gene with IGS, promoter and the transcribed region, from which many Pol I molecules elongate the pre-rRNA. The transcribed region (labeled in green) is composed of the 5' external transcribed spacer (ETS), 18S rRNA (labeled in dark green), internal transcribed spacer (ITS) 1, 5.8S rRNA (labeled in dark green), ITS2, 28S rRNA (labeled in dark green) and 3' ETS. Several

terminator repeats (labeled in red) are located 3' of the transcribed region. Spacer-promoter (labeled in yellow) and the actual promoter (labeled in yellow) are found in the 5' region of the gene. Enhancer repeats (labeled in blue) are located in between the two promoters. A scale bar for DNA length is shown on the right top corner of each schematic representation. **C** On top an electron micrograph showing a Miller spread is shown, which was originally published by Miller and Beatty, 1969[33] and represents a section of the micrograph shown in Potapova and Gerton, 2019[37]. The lower part shows a schematic representation of a 'Miller spread' depicting an actively transcribed gene. The rDNA gene is depicted as in (B) with the IGS (grey), the enhancer repeats (blue), the spacer and 47S rDNA promoters (yellow), and the transcribed region (green). Elongating Pol I molecules transcribing the 47S rDNA sequence synthesize the pre-rRNA (fibrils extending from the transcribed region). Ball-like structures represent co-transcriptional pre-ribosomal assembly intermediates.

1.3 Pol I structure and subunits

Pol I itself is a large protein complex of about 600 kDa, composed of 14 subunits in yeast[17]. Homologues of 13 subunits were identified in metazoans, but for subunit A14 no homologue was detected yet[17]. In mammals, about 10 % of Pol I extracted from the nucleus is in an initiation-competent state (Pol Iβ), whereas remaining Pol I (Pol Iα) represents elongating or 'to be converted to initiation-competent' Pol I which is not active in promoter-dependent *in vitro* transcription[17]. One apparent difference between the two states is the stable association of the conserved initiation factor RRN3 to Pol Iβ which is required to stabilize the initiation competent form of the enzyme[17].

The overall structure of Pol I resembles the bi-modular architecture of Pol II with a DNA-binding, central cleft, 'core' and 'shelf' modules and two additional protruding subcomplexes, the RPA49/34 heterodimer on the core module and the stalk at the shelf module[17,40–43]. The core polymerase is composed of the five common subunits shared between all three polymerases (RPABC1, RPABC2, RPABC3, RPABC4, and RPABC5), of two subunits shared with Pol III (RPAC1 and RPAC2), as well as three Pol I-specific subunits (RPA1, RPA2, and RPA12). RPA12 binds the RPA2 lobe-structure and interacts and anchors the two Pol I-specific subunits RPA49 and RPA34, building the protruding heterodimer. The stalk is constructed by the Pol I-specific subunit RPA43 and in yeast additionally by A14[17,42–44] (Figure 2A). Table 1 shows an overview about homologous Pol subunits in yeast Pol I and human Pols and their function.

The Pol I cleft (Figure 2B and C, green), into which DNA is loaded, is localized in between the clamp (Figure 2B and C, pink) and protrusion domain on the upstream edge (Figure 2B and C, dark blue), and between lobe (Figure 2B and C, blue) and jaw domains (Figure 2B and C, light blue) on the downstream edge. Next to the clamp, the stalk (Figure 2, slate) is bound to the core polymerase. In Pol I, clamp, stalk, and shelf module build a rigid mobile element, in which the shelf module consists of the cleft and foot domain of RPA1 as well as RPABC1 and RPABC2[36,45]. Protrusion, lobe and jaw on the other side of the cleft are ordered from the upstream to the downstream edge of the enzyme. Subunit RPA12 (Figure 2, dark orange) and the heterodimer RPA49/34 (Figure 2, light purple/light pink) on the opposite side of the polymerase compared to the stalk make their contacts to the core Pol I via the lobe domain. The funnel (Figure 2B and C, yellow-green) and pore (Figure 2B and C, yellow) are located underneath the lobe/jaw region and the cleft and enable NTPs and the C-terminal domain of RPA12 to reach the active center (Figure 2B and C, orange). The active center is located behind the cleft and is terminated by the wall (Figure 2B and C, dark green) which is located at the upstream side of the polymerase and helps binding the promoter DNA together with the protrusion[36].

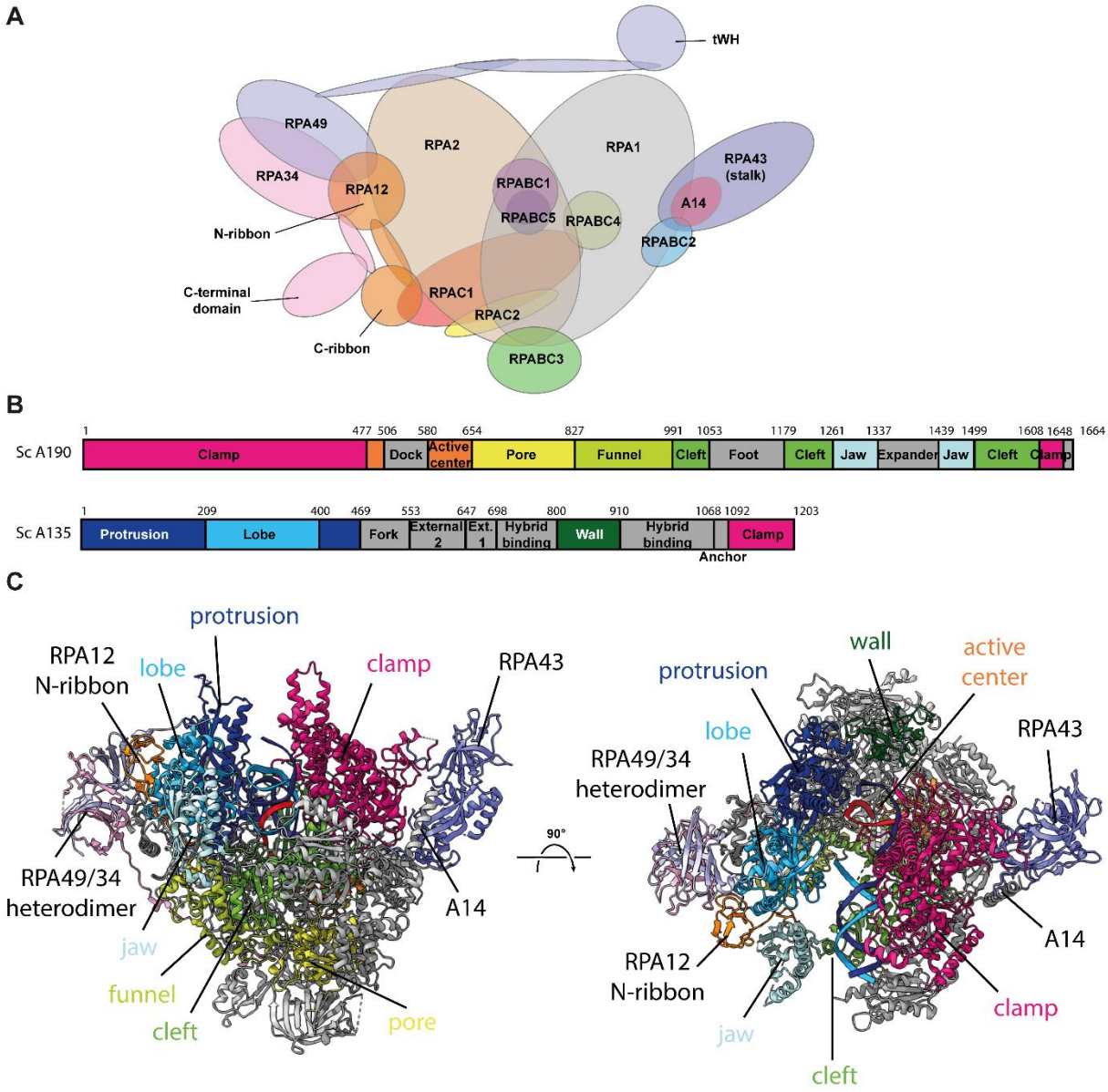


Figure 2 Architecture and functional sub-domains of yeast Pol I

A Schematic 2D representation of the Pol I structure with colored subunits (Pol I subunit colors are: RPA1 grey, RPA2 wheat, RPAC1 red, RPAC2 yellow, RPABC1 magenta, RPABC2 hafnium, RPABC3 green, RPABC4 lemon, RPABC5 density, RPA43 slate, RPA12 dark orange, RPA49 light blue, RPA34 light pink). **B** 2D schematic representation of yeast subunits A190 and A135 (to scale) including domain boundaries: clamp (pink), protrusion (dark blue), lobe (blue), jaw (light blue), cleft (green), funnel (yellow-green), pore (yellow), active center (orange) and wall (dark green) **C** Cartoon representation of elongating *S. cerevisiae* Pol I (code 5M3F)[36] outlining functional sub-domains in the 3D model (colored as in (B)) and subunits RPA49 (light purple), RPA34 (light pink), RPA12 (dark orange) and RPA43 (stalk; slate).

Table 1 Nomenclature and function of Pol I, (Pol II and Pol III) subunits in yeast and human

Proteins and genes in brackets are associated to the respective polymerase, but not a subunit of it.

S. cerevisiae Pol I		*H. sapiens* Pol I		*H. sapiens* Pol II		*H. sapiens* Pol III		Function
protein	***gene***	***protein***	***gene***	***protein***	***gene***	***protein***	***gene***	
A190	*RPA190*	RPA1	*POLR1A*	RPB1	*POLR2A*	RPC1	*POLR3A*	active center
A135	*RPA135*	RPA2	*POLR1B*	RPB2	*POLR2B*	RPC2	*POLR3B*	active center
AC40	*RPC40*	RPAC1	*POLR1C*	RPB3	*POLR2C*	RPAC1	*POLR1C*	
AC19	*RPC19*	RPAC2	*POLR1D*	RPB11	*POLR2J*	RPAC2	*POLR1D*	
Rpb5	*RPB5*	RPABC1	*POLR2E*	RPABC1	*POLR2E*	RPABC1	*POLR2E*	
Rpb6	*RPO26*	RPABC2	*POLR2F*	RPABC2	*POLR2F*	RPABC2	*POLR2F*	
Rpb8	*RPB8*	RPABC3	*POLR2H*	RPABC3	*POLR2H*	RPABC3	*POLR2H*	
Rpb12	*RPC10*	RPABC4	*POLR2K*	RPABC4	*POLR2K*	RPABC4	*POLR2K*	
Rpb10	*RPB10*	RPABC5	*POLR2L*	RPABC5	*POLR2L*	RPABC5	*POLR2L*	
A12.2	*RPA12*	RPA12	*POLR1H/ ZNDR1*	RPB9	*POLR2I*	RPC10	*POLR3K*	RNA cleavage, proof-reading, termination
				(TFIIS)	*(TCEA1)*			
A43	*RPA43*	RPA43	*POLR1F/ TWISTNB*	RPB7	*POLR2G*	RPC8	*POLR3H*	open complex formation, elongation, termination
A14	*RPA14*	-	-	RPB4	*POLR2D*	RPC9	*CRCP*	open complex formation, elongation, termination
A49	*RPA49*	RPA49	*POLR1E/ PAF53*	(TFIIFα)	*(GTF2F1)*	RPC5	*POLR3E*	initiation, elongation
				(TFIIEβ)	*(GTF2E2)*			
A34.5	*RPA34*	RPA34	*POLR1G/ CD3EAP, CAST, ASE1, PAF49*	(TFIIFβ)	*(GTF2F2)*	RPC4	*POLR3D*	initiation, elongation
						RPC3	*POLR3C*	open complex stabilization
						RPC6	*POLR3F*	open complex stabilization
						RPC7	*POLR3G*	open complex stabilization

The two largest subunits, RPA1 and RPA2, build the active center and are responsible for the enzymatic activity, such as substrate binding, RNA elongation, and template movement[17]. The DNA-mimicking 'expander' loop of subunit RPA1 becomes structured in inactive yeast conformations and resembles the DNA backbone location in the cleft, thus preventing unspecific DNA binding[42–44]. The bridge helix within RPA1 (Sc P992-L1027; Hs P961-L996) is involved in the nucleotide addition cycle and resembles a 'nanomechanical switchboard' required for physical translocation after successful phosphodiester bond formation as first shown the yeast Pol II homologue[46–48]. In monomeric states the RPA1 bridge helix can be unwound and becomes folded in elongating polymerases[5]. In yeast, Pol I dimerization is possible and resembles an inactive state, but is not exclusive to *S. cerevisiae* as previously speculated[42–44,49]. During activation of Pol I from dimers or monomers to the elongating state structural rearrange-

ments occur: specifically, relocation of the 'expander' to allow for DNA binding, and contraction of the active center cleft to allow folding of the bridge helix in order to enable RNA synthesis[5,42]. Such contraction of the active center cleft is accomplished by a movement of the clamp and stalk region and is specific to Pol I[5,36,44,50].

The stalk subcomplex is able to carry out multiple functions in all Pols across all domains of life including promotion of open complex formation during initiation, the increase of processivity, and augmentation of transcription termination[3].

Subunit RPA12 is important for proper termination and promotes RNA cleavage as well as proofreading activity of Pol I[51]. The C-terminal ribbon (C-ribbon) of RPA12 resembles the C-ribbon of the Pol II elongation factor TFIIS, which can enter the funnel to reach the newly synthesized RNA in a backtracked state, and hence enables the intrinsic cleavage activity of Pol I[42,43,52,53]. This C-ribbon is connected to an N-ribbon (N-terminal ribbon) domain, which is stably associated with the Pol I lobe. The linker between both RPA12 domains is flexible, allowing detachment of the C-ribbon and reaching the active center via the nucleotide triphosphate (NTP) entry pore domain to promote RNA cleavage[52,54]. The N-terminal domain of RPA12 resembles the N-terminal domain of Pol II subunit RPB9 and binds the lobe domain of Pol I subunit RPA2[54,55].

The heterodimer RPA49/34 functionally supports initiation and elongation stages[5,56,57]. The subunits interact via their N-terminal dimerization domains and contact the Pol I core at the RPA2 lobe and the N-ribbon of RPA12[58]. The RPA49/34 dimerization domain structurally and functionally resembles Pol II transcription factor TFIIF[52,56,59]. Subunit RPA34 also possesses a linker 'ARM' running along the core polymerase and a large flexibly linked, non-conserved C-terminal domain (CTD)[42,43]. RPA49 contains a linker including a helix-turn-helix (HTH) motif and a tandem winged helix (tWH) domain, comparable to Pol II transcription factor TFIIE[5,52,56,60,61]. Both HTH and tWH of RPA49 possess DNA binding activity with the tWH showing a sequence preference to promoter DNA[5,56,60]. Upon Pol I cleft contraction, the linker locates above the cleft and contacts the non-template (nt)-DNA strand during transcription initiation[62,63]. Detailed functional analysis of the yeast RPA49/34 heterodimer uncovered diverse functions: During initiation the tWH of RPA49 might help to recruit transcription initiation factor RRN3 and facilitates its release upon promoter escape. Additionally, tWH and HTH of RPA49 is supposed to help to orient Pol I, select transcription start site, and structurally rearrange Pol I into a catalytic active conformation[5,59,60]. During elongation the heterodimer enhances processivity, stimulates RNA cleavage activity, and facilitates the passage through nucleosomes[56,57,64]. The heterodimer may also support high loading rates of Pol I on one rDNA gene. While direct contact of adjacent Pol I molecules has been proposed in a head to tail manner via RPA49 and RPA43 interaction[65], such a mechanism is challenged as contacts cannot be detected in cryo-electron tomography analysis of actively elongating Pol I molecules in close-to-native conditions[36]. RPA49/34 is deemed to be a global center of Pol I regulation as the subcomplex can be dissociated under certain circumstances and can get lost during purifications. Accordingly, specific regulatory modifications, such as acetylation, have been observed in mammals[17,66,67]. Furthermore, the sub-complex occupancy varies among organisms and heterodimerization is species-specific and not possible across species[44,60,65,66,68].

1.4 Factors involved in Pol I transcription

Many additional factors assist the polymerase within cells to ensure proper Pol I transcription, among which the initiation factors are most extensively studied. In mammals, transcription factor RRN3, 'upstream-binding factor' (UBF) and 'selectivity factor 1' (SL1) enable proper transcription initiation. RRN3 is essentially conserved between yeast and human[11], whereas core promoter-binding factor SL1 includes the homologues of yeast core factor (CF) subunits, but additionally contains TBP (TATA-box-binding protein) and at least one additional subunit (TAF1D)[11,23,69–71] (see Figure 3 for domain organization and conservation of initiation factors). Yeast CF consists of subunits Rrn6 (homologue to human TAF1C), Rrn7 (TAF1B), and Rrn11 (TAF1A) in a 1:1:1 stoichiometry, binds the core promoter DNA, and directly recruits Pol I-Rrn3 complex[39,62,72,73]. The mammalian Pol I transcription initiation system also includes homodimeric UBF, which binds the UPE as well as the transcribed rDNA region. UBF is distantly related to the yeast Hmo1 protein, the UBF C-terminal domain (CTD) shows similarities to the C-terminal region (CTR) of Net1 in the yeast system, and may have partial functional conservation to the UPE-binding 'upstream activating factor' (UAF) in yeast[69,71,74,75]. In yeast, HMG (high mobility group) -box containing protein Hmo1 shows homology with UBF HMG boxes 1 and 2, and has functional similarities by binding throughout the rDNA gene and stabilizing the nucleosome depleted region (NDR)[76–80]. Hmo1 contains a poorly conserved HMG box A and a conserved HMG box B responsible for dimerization and DNA binding, respectively[79]. The CTR of yeast Net1 (aa 1052-1189) stimulates Pol I transcription *in vitro* and *in vivo*[81]. Yeast UAF contains subunits Rrn5, Rrn9, Rrn10, Uaf30, and histones H3 and H4[82,83]. UAF binds the UPE and thereby prevents rDNA transcription by Pol II, but so far has no mammalian counterpart[83,84]. UAF together with TBP increases transcription from basal levels 10-50 fold by bridging between CF and UAF via direct protein-protein interactions and likely independent from its known DNA-binding to TATA boxes[85–88].

RRN3

RRN3 association to Pol I leads to structural rearrangements[89,90] and is regulated by post-transcriptional modifications, including changes during the transcription cycle. But only a minor portion of Pol I is bound to RRN3 at the same time[11,17,61,89,91–94]. With its direct interaction with SL1 and Pol I including binding to RPA43[91,92,94,95], RRN3 plays a central role in Pol I recruitment to the promoter and bridging Pol I to promoter-bound initiation factors[17,95]. RRN3 stabilizes the monomeric, initiation competent Pol I. Thus, it promotes formation of the pre-initiation complex and the formation of the first phosphodiester bonds[61,90,91], making the factor essential during initiation[89,96,97]. Interesting is also the fact that human RRN3 is capable to interact directly with DNA *in vitro*[98].

SL1

Initiation factor SL1 is composed of the three core factor homologues TAF1A (Rrn11), TAF1B (Rrn7), TAF1C (Rrn6), and stably includes TBP and TAF1D[17,70,99,100]. TAF12 was proposed to constitute an additional subunit of the SL1 complex *in vivo*[101]. SL1 weakly interacts with RPA34[102], whereas it can interact directly with RRN3 via its subunits TAF1B and TAF1C[17,69,91,103]. SL1 is essential for mammalian Pol I transcription[17,70,104] and directs its species-specificity, hence its name 'selectivity factor 1'[17,70,100,105]. For instance, Pol I and UBF are interchangeable between mice and human, but SL1 is not and only

promotes rDNA transcription of the rDNA gene from its original source organism[70,106]. Overall, SL1 functions in promoter recognition and binding of the core promoter element and then recruits RRN3-bound Pol I to the promoter[17,103,107]. However, stable association of SL1 to the promoter requires UBF, otherwise SL1 cannot associate efficiently[14,106]. Without SL1, UBF frequently de- and re-associates from/to promoter DNA, but after binding of SL1, both factors remain stably bound over many rounds of transcription initiation[17]. Further supporting functional cooperativity of SL1 and UBF binding to the promoter, DNase I foot-printing was shown to be in line with mutual stabilization of the factors[17,108].

UBF

UBF assembles as a homodimer[108,109], specifically binds to the UPE and the core promoter, and functions as an architectural protein at the rDNA promoter[17,108]. UBF additionally binds throughout the transcribed rDNA region, defining the structure of actively transcribed genes[19,110,111]. At the same time UBF replaces core histones and nucleosomes on active rDNA genes, creating a nucleosome depleted region, and plays a role in regulating Pol I elongation[18,19,112]. UBF is composed of an N-terminal dimerization domain responsible for homodimerization, followed by six consecutive HMG-box domains and a highly charged, acidic CTD[11,17,108,109,113]. HMG-boxes 1-3 function in DNA binding which induces bending of an about 140 bp long DNA segment to a 360° loop[17,109,114–116]. HMG box 1 appears to be necessary and sufficient for sequence-specific promoter binding[108]. Binding and recruiting SL1 occurs via the CTD of UBF in a phosphorylation-dependent manner and might involve TBP and TAF1A of SL1[11,17,117–119]. The interaction between UBF and TBP relies on the conserved C-terminal region of TBP[117]. Promoter-bound UBF functions during Pol I promoter escape at the transition between initiation and elongation[120] and can directly interact with the Pol I heterodimer subunits RPA49 and RPA34[102]. Interestingly, UBF is expressed in two splice variants, the 'full-length' UBF1 and UBF2 lacking a part of the DNA-binding HMG box 2[14,108,121]. Both, UBF1 and UBF2, bind throughout the rDNA gene in the NDR, but only UBF1 appears to be bound at the spacer and gene promoters[18]. This observation is astonishing as it was shown that HMG box 1 is important and sufficient for sequence specificity[108] and although box 1 is present in both splice variants, the two variants show different localizations. While the reason of this remains unclear, cooperativity with SL1 may play a role in the process.

Additional factors

Besides the above factors, numerous other proteins participate in Pol I transcription and its regulation. One of these factors, 'transcription termination factor I' (TTF-I) is a myb-domain containing factor related to the yeast proteins Reb1 and Nsi1 and binds to the terminator site of human rDNA genes. As mentioned before, TTF-I has additional binding sites within the promoter region (Figure 1B) and this constellation is conserved in yeast. Besides its role in termination, TTF-I may additionally act as a replication fork barrier[122,123], or may activate and silence transcription[124–127]. Another factor necessary for termination is 'Pol I transcript release factor' (PTRF), helping to release Pol I from the DNA and afterwards allowing Pol I to be primed for re-initiation by re-association of RRN3 or to be stored[128,129]. Transcription factor PAF1C, helping during Pol I transcription elongation *in vivo*, is a transcription factor well known from the Pol II transcription system[130]. The histone chaperone 'facilitates chromatin transcription' (FACT) is also associated with the rDNA gene and plays an important role in Pol I transcription elongation through nucleosomes[131].

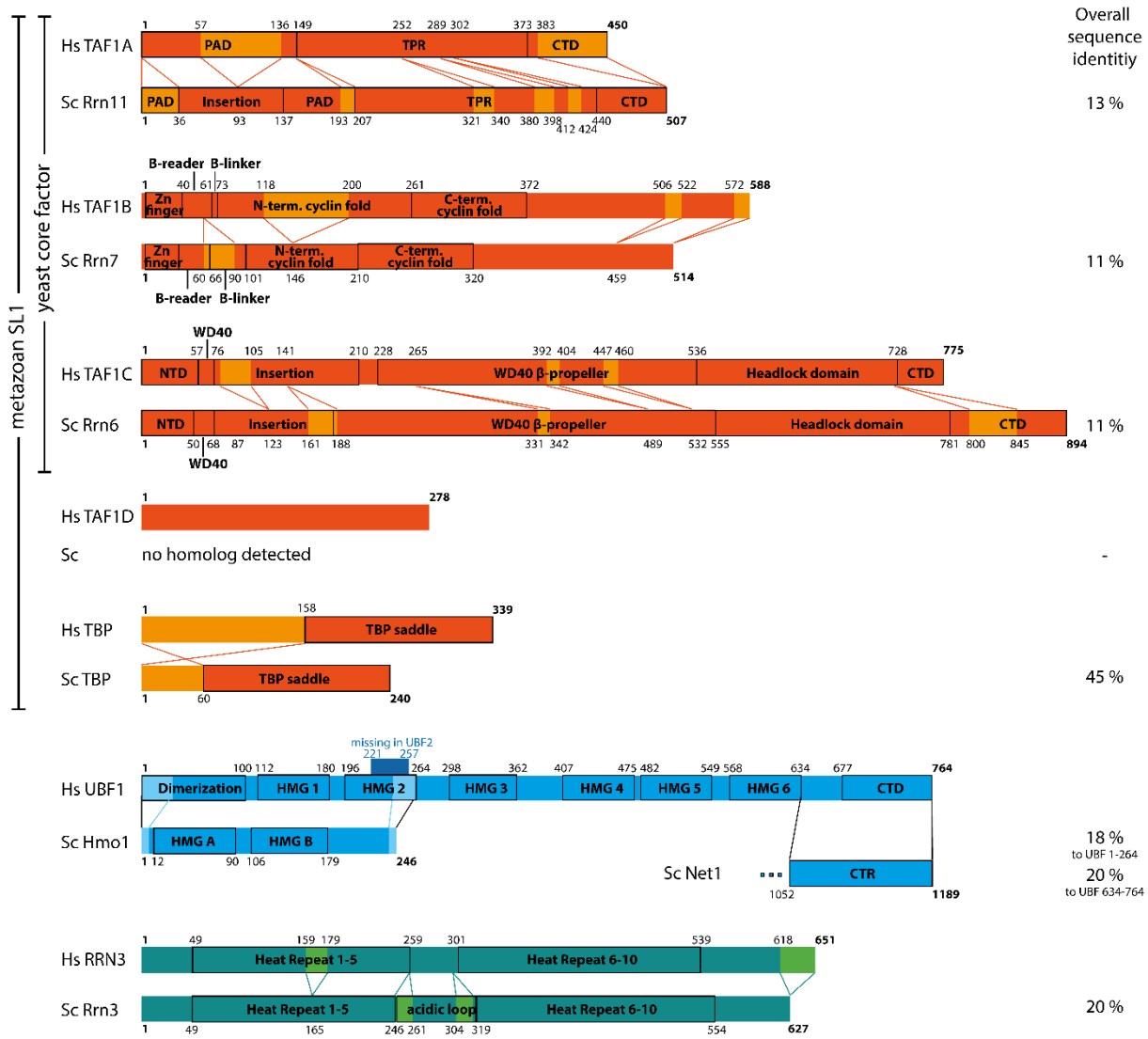


Figure 3 Pol I transcription initiation factors

Schematic domain architecture of human (Hs) Pol I transcription initiation factors with largest differences to their yeast (Sc) homologues: TAF1A (Rrn11), TAF1B (Rrn7), TAF1C (isoform 2; Rrn6), TAF1D, TBP (TBP), UBF (Hmo1/Net1) and RRN3 (Rrn3). SL1 subunits are shown in orange, UBF in blue and RRN3 in teal. Subdomains and insertions/deletions of 10 or more residues are indicated.

1.5 Pol I transcription cycle

At actively transcribed, open rDNA genes the promoter is bound by Pol I specific transcription initiation factors (SL1 and UBF1), which help to recruit RRN3-bound Pol I molecules. After successful transcription initiation and promoter escape, Pol I proceeds into productive elongation phase. Initiation factors UBF and SL1, but not RRN3, remain bound to the promoter during Pol I promoter escape. This stable binding allows fast recruitment of the next Pol I-RRN3 complex and thereby promotes transcription initiation. During the transition into the productive elongation phase, RRN3 also dissociates from Pol I. In case of pausing, stalling, or incorrect nucleotide incorporation into the nascent RNA during elongation, the built-in cleavage and backtracking activity of Pol I allows re-entering into productive

elongation. After transcription termination, RRN3 can re-associate with released Pol I to allow for re-initiation at the promoter.

For the yeast enzyme, detailed analysis of the transcription cycle was performed and for most of the steps various structures were obtained[36,42,43,50,61,90,132,133] (Figure 4). The mechanism of 'hibernation by dimerization' is found in yeast and is conserved from *S. cerevisiae* to *S. pombe*[5,42–44,49,134–136]. Pol I dimers are in a concentration- and salt-dependent equilibrium with monomers and trapped in an inactive state for transcription[42,44]. Monomeric Pol I gets stabilized by the association of initiation factor Rrn3 and primed for initiation[89,90]. The association of Rrn3 is phosphorylation-dependent and dissociates from Pol I during promoter escape[94]. In yeast, the three-subunit core factor binds the core promoter, the six-subunit UAF binds the UPE, and TBP bridges between the two protein complexes[5]. Promoters bound by its transcription initiation factors can recruit Rrn3-bound Pol I and enable transcription initiation, which was extensively studied structurally[39,62,63,132,133]. During transcription initiation, structural re-arrangements of Pol I and core factor take place. The contraction of the Pol I active center cleft assists in DNA melting and stabilization of the opened transcription bubble. After promoter clearance and loss of Rrn3, Pol I enters productive elongation phase. Cleft contraction is a gradual mechanism from the dimer/monomer to elongation complexes[5], possessing a fully contracted active center cleft during elongation[36,44,50,55]. The catalytic nucleotide addition includes selection of the correct nucleotide, its addition to the 3' end of the nascent RNA, and release of the cleaved pyro-phosphate before translocation to the next nucleotide[137]. In contrast to Pol II, Pol I is able to transcribe through nucleosomal templates[64,131] and is about 1.5 times faster and pauses less often than Pol II[138]. During pausing, backtracking of the polymerase is a common consequence, which means that the enzyme diffuses in the retrograde direction on the DNA strand[3,138]. During backtracking the nascent RNA is extruded through the nucleotide entry pore and thereby prevents Pol I from further elongation[3]. To overcome backtracking, the nascent RNA is cleaved, creating a new 3' end at the RNA in the active center[3,138]. Yeast Pol I subunit A12.2 (RPA12 in humans) harbors the catalytic site responsible for cleavage of the nascent RNA and is required for cleavage and termination activity of Pol I[138,139]. In yeast, Pol I termination factor Nsi1 binds to its binding sites at the terminator and together with a T-rich stretch, this 'road block' seems to be sufficient for transcription termination *in vitro*[128,140,141]. In mammals, the homologue TTF-I binds a conserved 18 bp DNA sequence and stops Pol I *in vitro*[128,142,143]. PTRF is additionally needed in mammals to release Pol I from the DNA and afterwards allows Pol I to be primed for re-initiation by re-association of RRN3 or to be stored[128,129].

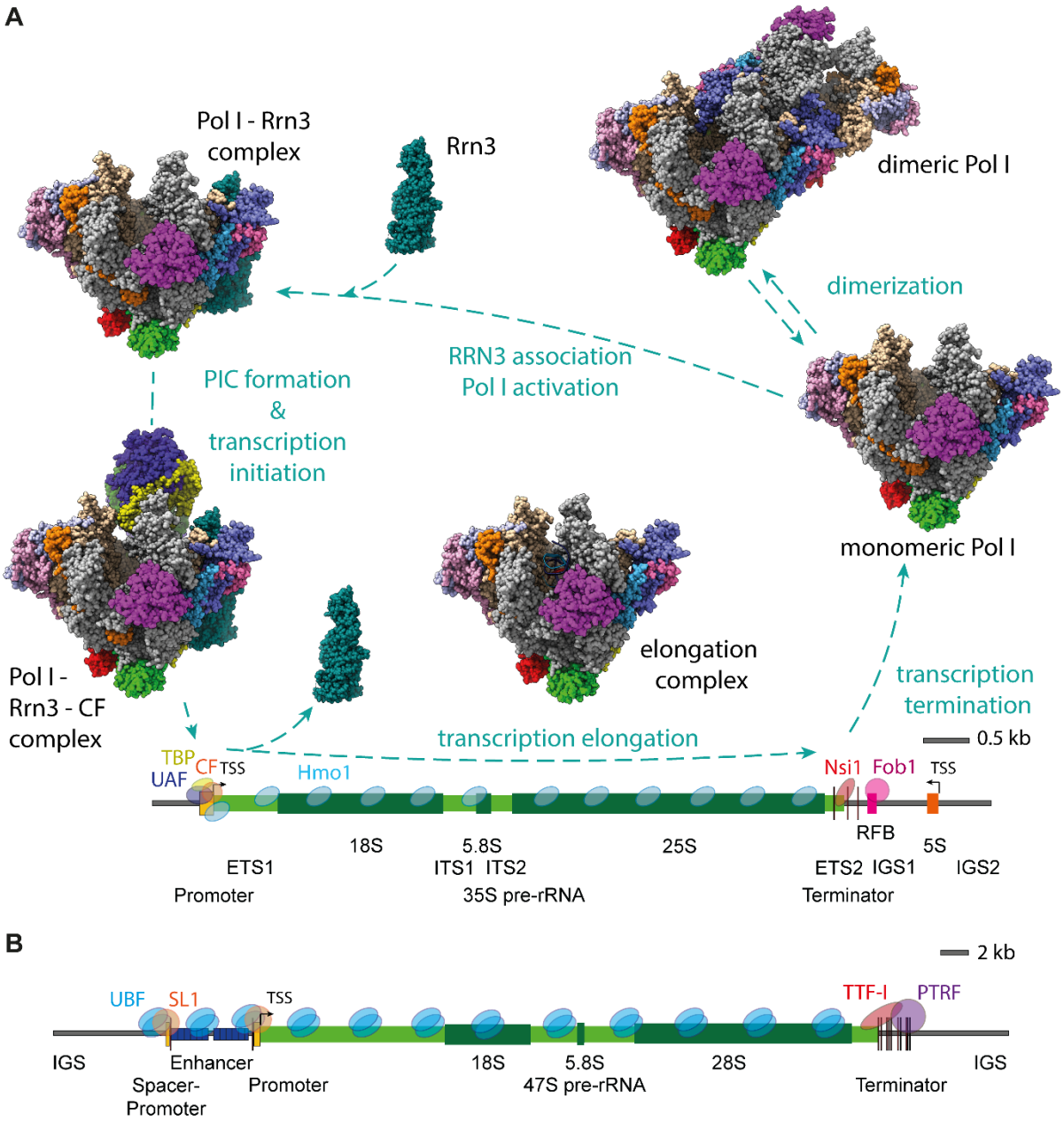


Figure 4 Yeast Pol I transcription cycle

A Monomeric yeast Pol I (colored by subunits as in Figure 2A) can be activated by Rrn3 (teal) association and primed for initiation. Promoter-bound UAF (purple), TBP (yellow-green) and CF (orange) recruit the Pol I-Rrn3 complex and transcription initiation takes place. After promoter escape and Rrn3 dissociation, Pol I enters productive transcription elongation. HMG-box protein Hmo1 (light blue) is binding throughout active genes. Terminator-bound factor Nsi1 (red) is essential for transcription termination and the replication fork barrier (RFB) (pink) is bound by Fob1 (pink). After termination, monomeric Pol I can re-enter into the cycle. Monomeric yeast Pol I can also be stored as dimers following the mechanism 'hibernation by dimerization'. Bottom: Schematic representation of the yeast rDNA genomic locus comparable to Figure 1 (5S rRNA gene is shown in orange). **B** As comparison the human rDNA gene locus is shown as in (A) and Figure 1 and additional transcription factors are introduced at their binding sites: UBF (blue), SL1 (orange), TTF-I (red) and PTRF (purple). (PDB accession codes: 5G5L (Pol I-Rrn3 complex and Rrn3), 5N60 (Pol I-Rrn3-CF complex), 5M3F (EC), 5M3M (monomer), 4C2M (dimer); Pol I subunit colors are: A190 grey, A135 wheat, AC40 red, AC19 yellow, Rpb5 magenta, Rpb6 hafnium, Rpb8 green, Rpb12 lemon, Rpb10 density, A43 slate, A12.2 orange, A49 light blue, A34.5 light pink, Rrn3 teal, Rrn6 purple, Rrn7 green, Rrn11 yellow. Colors of schematic rDNA locus: IGS (grey), Transcription start site (arrow), promoter (yellow), enhancer (blue) transcribed region (green), 18S, 5.8S, and 25S rRNAs (dark green), terminator (red), replication fork barrier (pink), 5S rRNA gene (orange))

1.6 Regulation of Pol I transcription

Considering the importance of Pol I transcription with its direct link to cell growth and cell division, the transcription system must be tightly regulated[17]. One regulatory strategy is the 'opening' (activation) and 'closing' (inactivation) of single rDNA genes. However, this mechanism was shown to play a minor role in regulation[82]. Even in fast growing cells only about 50 % of genes are transcriptionally active and in an open chromatin state. Whereas Pol I transcription drops by about 90 % in stationary phase cells, opened genes are only reduced about half[82]. Under growth stimulation Pol I transcription increases several-fold, but the number of active genes stays constant[11,144]. Furthermore, phenotypic analysis of yeast strains carrying a reduced number of rDNA repeats are viable without any apparent growth defect under normal conditions while all repeats remain open and constitutively activated[145]. Generally epigenetic control of repeat activation may have more importance in mammalian cells[146], in yeast the presence of multiple rDNA gene copies appears to be required for genome stability and control repeat amplification[147], rather than transcription boosting[148].

Post-translational modifications (PTM) of proteins involved in Pol I transcription are another possibility of regulation. Initiation factors RRN3 and UBF as well as Pol I subunits RPA1, RPA2 and RPA49/34 are phosphoproteins[68,149–155] and therefore may serve as regulation hubs[11,17,102]. For example, the 'protein kinase C' (PKC) pathway, which monitors the cell surface integrity, is able to repress Pol I transcription in case of secretory defects in yeast[82,156]. Moreover, phosphorylation of the CTD of UBF by 'casein kinase 2' (CK2) leads to an increase of UBF-SL1 interaction and thereby contributes to Pol I transcription activation and adjustment to cellular growth[11,17,108,118,157]. Overall, UBF activity and PIC (pre-initiation complex) formation is regulated by the growth-dependent 'mitogen-activated protein kinase' (MAPK) and 'extracellular signal-regulated kinase' (ERK) pathways as well as the growth-dependent and nutrient-sensing 'mammalian target of rapamycin' (mTOR) and stress dependent 'Jun N-terminal kinase 2' (JNK2) pathway[17]. Furthermore, regulation of Pol I association of RRN3 in a phosphorylation-dependent manner is an important regulatory strategy allowing the cell to rapidly react to changing nutrient availability and other stress conditions via a combination of 'ribosomal S6 kinase' (RSK), ERK, and mTOR pathways as well as the JNK pathway with enhanced repression of rRNA synthesis[11,17,82,92,158–160]. During nutrient or energy shortage, phosphorylation status of RRN3 is changed in a way that avoids interaction of RRN3 with promoter-bound SL1 by various kinases and phosphatases including 'AMP-activated protein kinase' (AMPK)[67,159,161]. Acetylation of RPA49 is another point of transcription regulation, which is performed by the interplay of 'CREB-binding protein' (CBP) and 'NAD-dependent protein deacetylase sirtuin-7' (SIRT7)[67]. SIRT7 is able to catalyze subunit deacetylation leading to increased Pol I transcription and is regulated in the step of SIRT7 nucleolus-localization in response to different stress conditions[67]. In fact, upregulation of SIRT7 expression was associated with several cancers, emphasizing its important role in cell growth regulation[67,162]. Another level of regulation might be the association of the heterodimer RPA49/34 to Pol I, which at least seems to be species-specific[66]. Although many of the known influences of regulation include PTM of initiation factors, regulation of transcription elongation seems to play an additional role[11,112] and some factors may prevent complete pre-rRNA synthesis by blocking elongation in a road-block-like fashion[163]. Figure 5 visualizes the regulatory network of human Pol I activity in a composite manner.

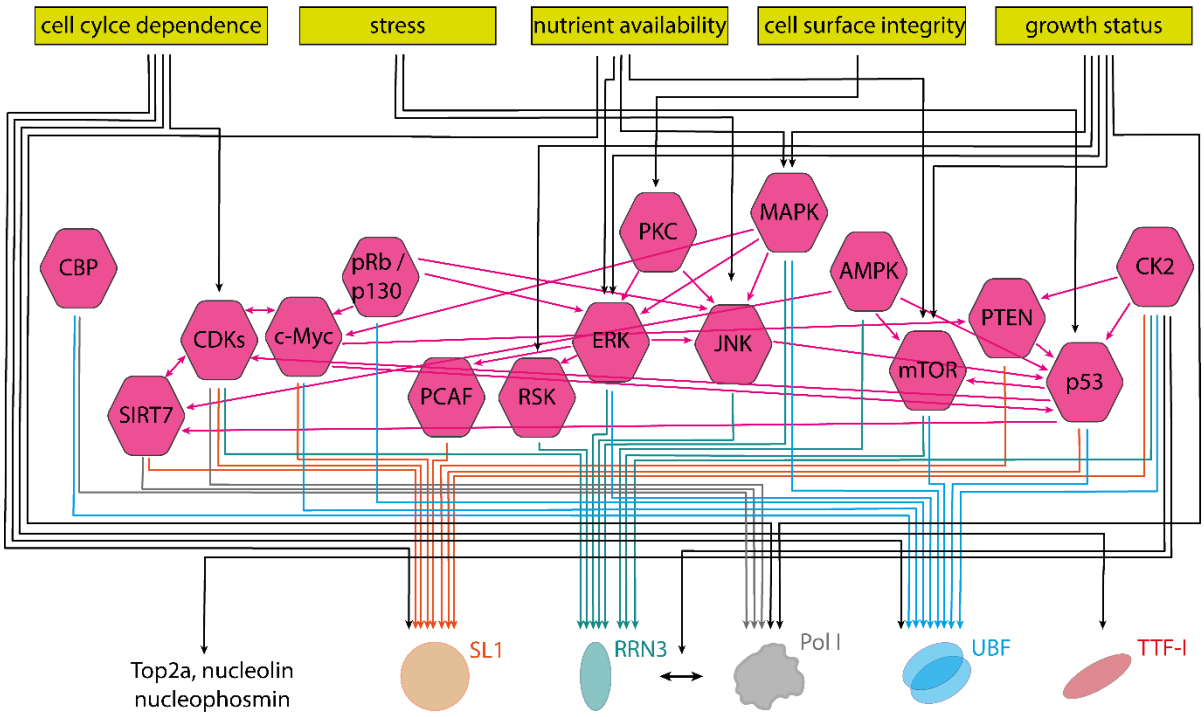


Figure 5 Schematic overview of the Pol I regulation network

Schematic overview of the regulation network of Pol I transcription: different cellular conditions (yellow) activate diverse signaling pathways and proteins (pink) that regulate one or more factors involved in Pol I transcription (arrow in the color of the regulated protein: SL1 (orange), RRN3 (teal), Pol I (grey), UBF (light blue), TTF-I (red)). Cross-talk between the different signaling pathways is indicated with arrows (pink). Regulation network was summarized using various references[17,150,158,159,164–167].

In many cancer cells, Pol I transcription is upregulated to cater the increased requirement for ribosomes in rapidly growing cells. Nevertheless, the underlying molecular mechanisms are diverse and remain elusive[72]. For example, the well-known 'Myc proto-oncogene protein' was shown to control Pol I transcription, which could represent one early step in oncogenesis in some cases[17,168,169]. Additionally, it was shown that several protein kinases have the potential to upregulate rDNA transcription in cancer cells, including CK2 and ERK, which lead to UBF and RRN3 phosphorylation[105,116,150,170,171]. Furthermore, mutations in UBF may underly some forms of acute myeloid leukemia (AML) and related pathologies[172,173]. Taking these facts into account, drugs downregulating Pol I transcription are potentially promising anticancer drugs, which started to be evaluated and will be further investigated[72]. In fact, targeting of initiation complex formation[174] or prevention of promoter escape[175] appear to be promising strategies.

For further information on yeast and human Pol I transcription the review of Daiß *et al.*[176] summarizes the current knowledge about this sophisticated transcription system in more detail and is based on and contains large parts of the introduction, as well as parts of the discussion of this thesis.

1.7 Scope of this study

Transcription of DNA is a fundamental mechanism of each living cell and is carried out by at least three nuclear Pols[2–4]. Pol I recognizes only one specific promoter, transcribing exclusively the 47S pre-rRNA gene[10–12]. The four rRNAs account for about 80 % of total RNA within the cell[11,18]. Over the last years, Pol I transcription came into focus of modern medicine, as an increasing number of cancer types are associated with mutations within the Pol I transcription machinery[18,22]. Different drugs downregulating Pol I transcription are investigated as anti-cancer-treatment strategies[165,167]. Nevertheless, Pol I transcription has been mostly studied in lower model organisms due to experimental limitations[6,45]. Yeast Pol I transcription was examined in detail on a mechanistic, structural, and functional level, using purified and recombinantly expressed components. As a result, a clear definition of subunit (sub-) domain and transcription factor functionalities in transcription initiation, elongation, transcript cleavage, backtrack recovery and termination was achieved[36,42,43,50,51,61,64,90,132,133]. Contrary, human Pol I (hPol I) transcription has been mostly studied in extracts or partially purified systems[166,177,178]. Thus, it remains unclear whether the results of structure-function studies performed in yeast can be easily transferred to higher organisms. Additionally, in higher eukaryotes, regulatory variants dependent on tissue type, developmental state and cell-cycle stage add additional layers of complexity. This leads to the unanswered question how Pol I structurally and functionally adapted to the increased regulatory demands in human cells.

In this study, I aim to introduce a well-defined human *in vitro* system consisting of purified components. A human cell line was modified, so that an endogenous tag was added on subunit RPA1, in order to allow for purification of the multi-subunit hPol I. The tag was then used to purify hPol I to high purity and quality in a close-to-native state. A detailed structure-function analysis of purified hPol I was performed elucidating which features of Pol I transcription are conserved among eukaryotes, which features are human-specific and how they may impact the transcription machinery. A detailed phylogenetic analysis going along with the comparison of *in vitro* transcription of yeast and human Pol I was performed. Furthermore, the cryo-EM structure of hPol I was determined, and uncovered human-specific features were investigated in further detail. The structure of hPol I allowed mapping and evaluating the effects of different disease-associated mutations of different, mainly neurodegenerative diseases and could help to understand the underlying molecular mechanisms of these Pol I-associated diseases, potentially helping to improve therapy or drug research. The basic research on human Pol I conducted in this thesis contains a detailed structure-function analysis including the comparison to its yeast counterpart and lays the groundwork for further research on the human Pol I transcription system.

2 Results

2.1 HeLa-POLR1A cell line

2.1.1 Establishing a cell line with sfGFP-tagged hPol I by CRISPR/Cas9

The structure-function analysis requires *in vitro* purified compounds. Thus, a human cell line was established carrying specifically tagged human Pol I (hPol I) allowing for purification of endogenous hPol I and *in vitro* investigations. In particular, hPol I subunit RPA1 was tagged with a C-terminal sfGFP (superfolder green fluorescent protein). For cell line establishment the same strategy as for tagging the hPol I/III shared subunit RPAC1[179] was used to tag Pol I-specific subunit RPA1 and the protocol was thereby proven to be reliable in general for homozygous knock-in of C-terminal fusion tags. For genome editing, the HeLa P2 cell line[180] was modified to the specific interests using the CRISPR/Cas9 system (see method section 6.5.7). The double-nickase strategy was applied, in which two different guide RNAs (gRNAs) target Cas9-nickase (Cas9n) to the two different DNA strands in close proximity to each other in order to minimize off-target effects. The resulting double strand break (DSB) close to the 3' end of the *POLR1A* gene encoding for subunit RPA1 was repaired with the help of introduced template DNA and the cellular 'homology-directed repair' (HDR) pathway. Thereby, the DNA sequence of a GS-linker containing a protease cleavage site and the open reading frame (ORF) of sfGFP was introduced just 5' of the endogenous stop codon of *POLR1A* (Figure 6A). Modified cells were individualized and sorted using 'fluorescence-activated cell sorting' (FACS). GFP-positive cells were expanded and analyzed for homozygous insertion. The results of the Western blot against the modified subunit RPA1 proofed the newly established cell line to be homozygous for the tag insertion by showing a single band of higher molecular weight compared to the parental cell line (Figure 6B). Additionally, PCR (polymerase chain reaction) on genomic DNA (gDNA) was performed with primers binding just outside the introduced sequence. PCR products from the parental cell line were specifically shorter than products from edited gDNA. The newly established cell line HeLa-POLR1A solely showed longer PCR products, which confirmed homozygous tag insertion (Figure 6C). The PCR products of the HeLa-POLR1A cell line was sequenced and the correct sequence verified.

2.1.2 sfGFP-tagged hPol I localized correctly in the nucleolus

Endogenous Pol I localizes to a nuclear sub-compartment the nucleolus, in which Pol I transcription takes place[10]. Complex integrity and correct cellular localization of tagged hPol I in the newly established cell line HeLa-POLR1A was examined to test whether the sfGFP-tag is really neutral or harmful for the enzyme. HeLa-POLR1A cells were grown, fixed, and imaged by confocal microscopy (see method sections 6.5.10 and 6.13). GFP signal, representing Pol I, was clearly detected, resembling many single dots within each nucleolus (Figure 6D). Each single dot could represent one active rDNA gene, on which many hPol I enzymes were transcribing. Thus, tagged RPA1 was correctly assembled within hPol I and the whole protein complex localized efficiently to the nucleolus, which is its place of action during interphase[10,181].

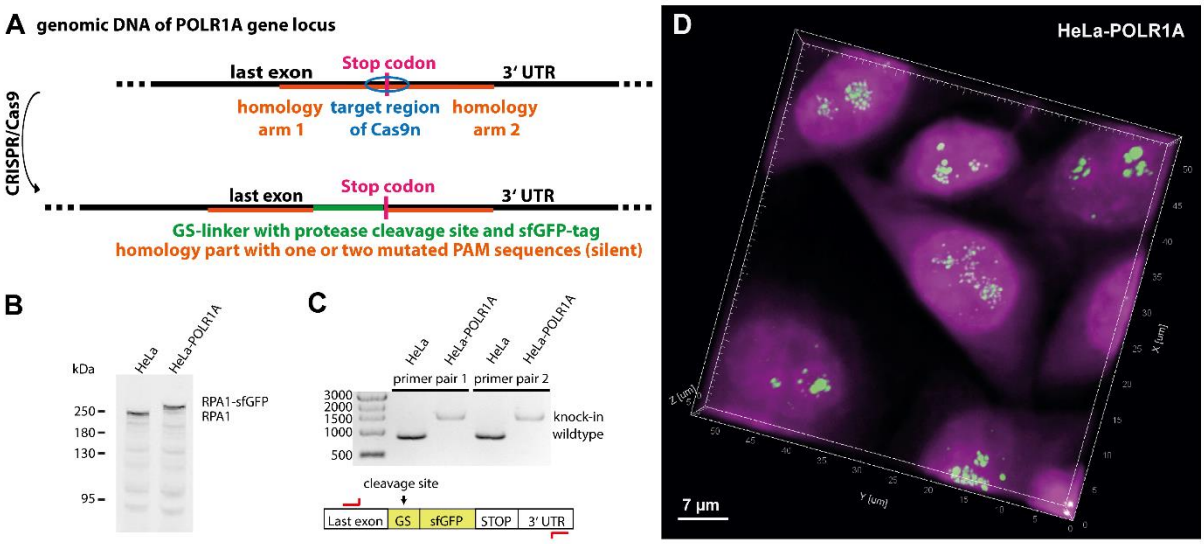


Figure 6 Human cell line HeLa-POLR1A

A The schematic overview of the used CRISPR/Cas9 strategy for the introduction of a protease-cleavable sfGFP tag (green) just 5' of the endogenous stop codon (pink) is shown. The target region of Cas9n cleavage (blue) is in the close surrounding of the stop codon and HDR is enabled by transfected template DNA containing the two homology arms (orange) and the tag sequence (green). **B** Western blot against subunit RPA1 showed a shift to higher molecular weight in lysates of the HeLa-POLR1A cell line, confirming exclusive expression of the modified protein. **C** Sybr-Safe stained agarose gel showed shifted bands of longer PCR products for HeLa-POLR1A cell line compared to the parental HeLa cell line. **D** Confocal imaging showed the exclusive location of GFP-induced fluorescence in the nucleoli in aligned 3D stacks. Spots in the central cell may represent single rDNA genes. Magenta: DAPI stain; Green: sfGFP signal (fused to RPA1).

2.1.3 Visually tracking hPol I localization during cell cycle

During mitosis and chromosome condensation the nucleolus temporary disassembles[182]. One not fully answered question in this process of mitosis is how an even distribution of Pol I enzymes is guaranteed between the two daughter cells. Our newly established cell line allowed us to track hPol I over the cell cycle. HeLa-POLR1A cells were roughly synchronized, fixed during mitosis, and imaged using confocal microscopy (Figure 7, see method sections 6.5.9, 6.5.10, and 6.13). During chromatin condensation in prophase, nucleoli and hPol I foci slowly dissolved while still having had some gathered hPol I in brighter dots. In prometaphase, hPol I essentially dissociated from the condensed DNA and spread out. Even distribution of hPol I over the cell was achieved in metaphase and lasted during ana- and in parts during telophase ensuring equal amounts in both daughter cells. Clustering of hPol I and reassembling of the nucleolus structure started in late telophase or early cytokinesis and allowed full functionality right after cell division. The obtained results are comparable with another study showing only a small portion of hPol I travelling with condensed chromosomes during mitosis using super-resolution microscopy[183].

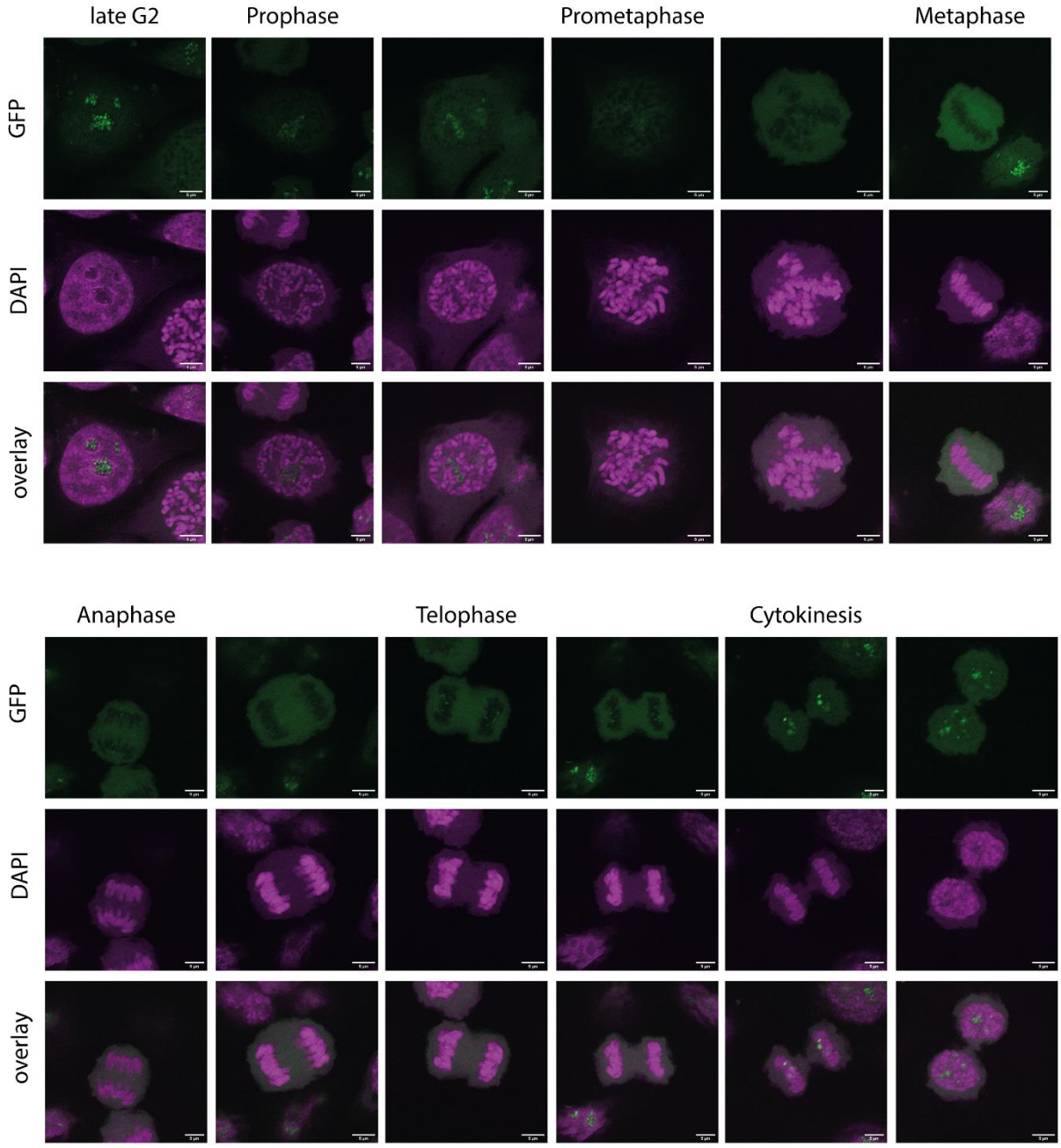


Figure 7 Tracking hPol I localization over the cell cycle by confocal microscopy

Confocal imaging shows the GFP-induced fluorescence resembling hPol I localization. GFP-signal (green) is shown in the top panel, below that the DAPI stain (magenta), and in the bottom channel the overlay of both channels. Confocal images were arranged following the cell cycle and the different phases are marked following cell cycle progression. Before mitosis, hPol I localized within single dots in the nucleolus. During cell division, hPol I signal spread out and hPol I was evenly distributed before the nucleolus with its hPol I dots started to re-appear during late telophase or cytokinesis. Scale bars: 5 µm

2.2 Purification of hPol I

Structure-function studies rely on highly purified *in vitro* proteins. Partially purified hPol I including additional factors already allowed some functional and biochemical investigations[17,103]. The possibility of affinity purification of hPol I now enables *in vitro* purifications to high purity allowing detailed functional and structural analysis. The sfGFP-tag on endogenous hPol I in the HeLa-POLR1A cell line was used for this purpose and a purification protocol implemented (see method section 6.7). Firstly, large scale suspension cell cultivation was established, followed by cell harvest and lysis. A GFP-pulldown was performed with the lysate and the captured hPol I was eluted from the beads by protease cleavage of the GFP-epitope (Figure 8A). After the GFP-pulldown, the complete 13-subunit Pol I enzyme was eluted (Figure 8B) and stochiometric amounts of heterodimer RPA49/34 were associated as judged from mass spectrometric analysis (Figure 9). The transcription initiation factor RRN3 co-purified with a subpopulation of hPol I enzymes (Figure 8B). Optionally, anion-exchange chromatography was performed to remove the protease within the sample. During this subsequent chromatography step 3C protease got separated, RRN3 dissociated, as well as large portions of heterodimer RPA49/34 (Figure 8C). Consequently, experiments were performed with the complete hPol I of the GFP eluate unless otherwise stated. In comparison, *S. cerevisiae* and *S. pombe* Pol I (Sc Pol I and Sp Pol I, respectively) were stable during anion-exchange chromatography and purified as complete 14-subunit complexes including its heterodimer[184]. Additionally, *S. cerevisiae* Pol I ΔA49/34.5 lacking the heterodimer A49/34.5 was used for functional studies (Figure 8D).

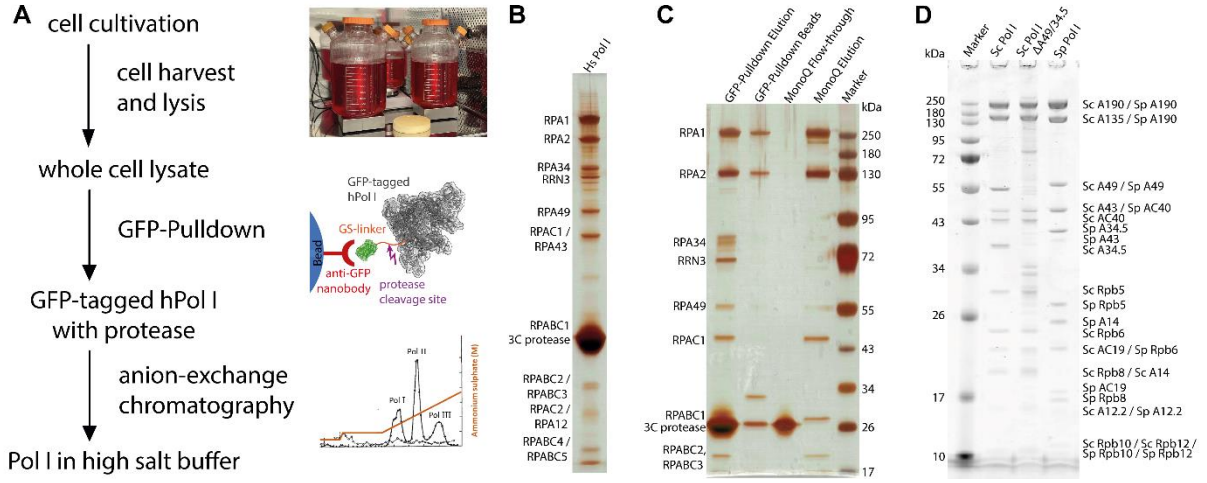
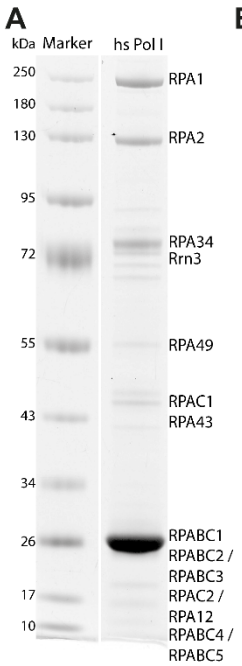


Figure 8 Purification of Pol I

A A schematic representation of the hPol I purification protocol is shown. (PDB 1GFL for GFP, 5m3m for Pol I; anion-exchange chromatography from Roeder and Rutter, 1969[2]) **B** Silver-stained SDS-PAGE of human Pol I GFP eluate showed bands for all subunits. **C** Silver-stained SDS-PAGE showed loss of initiation factor RRN3 and of large portions of the heterodimer RPA49/34 during anion-exchange chromatography: before anion-exchange chromatography all subunits were present, in the flow-through the 3C protease was found and for the hPol I elution fractions the bands of RRN3, RPA49 and RPA34 showed no or weak intensities. **D** Coomassie-stained SDS-PAGE of purified yeast Pol I variants showed bands for all subunits expected (Sc Pol I ΔA49/34.5 was missing the heterodimer A49/34.5 as expected).

The endogenously purified hPol I was examined concerning its composition using mass spectrometry (see method section 6.14.2). All known 13 human subunits as well as the transcription initiation factor RRN3 were identified with high confidence within the GFP-pulldown eluate. No other proteins possessed a sequence coverage over 25 % (Figure 9) and were identified with a similar confidence *in situ*. The yeast Pol I complexes are composed of 14 subunits, 13 of which are homologues to the mass spectrometry-detected human Pol I subunits. Yeast subunit A14 was only found in yeast Pol I and no comparable nor any protein was detected by mass spectrometry of purified hPol I besides the 13 described Pol I subunits and RRN3. Previously, it was already postulated that hPol I might lack yeast subunit A14 as no human homologue has been found using DNA and protein sequence-based searches[17]. Loss of the second stalk subunit A14 seemed indeed possible as depletion of *S. cerevisiae* A14 was not lethal, but resulted in a conditional growth phenotype[137,185], similar to what was reported from *S. pombe*[149]. Results from hPol I purification and mass spectrometry analysis suggest that hPol I is a 13-subunit enzyme.



B

subunit	score	No. of peptides	sequence coverage
RPA1	18730	233	75.7 %
RPA2	6765	106	58.9 %
RPAC1	3122	35	65.3 %
RPAC2	1604	22	66.2 %
RPA12	957	8	36.5 %
RPABC1	868	9	50.0 %
RPABC2	425	7	32.3 %
RPABC3	909	11	63.3 %
RPABC4	254	4	44.8 %
RPABC5	278	6	73.1 %
RPA43	564	9	29.0 %
RPA49	2525	39	53.2 %
RPA34	3770	50	71.2 %
RRN3	580	13	25.3 %

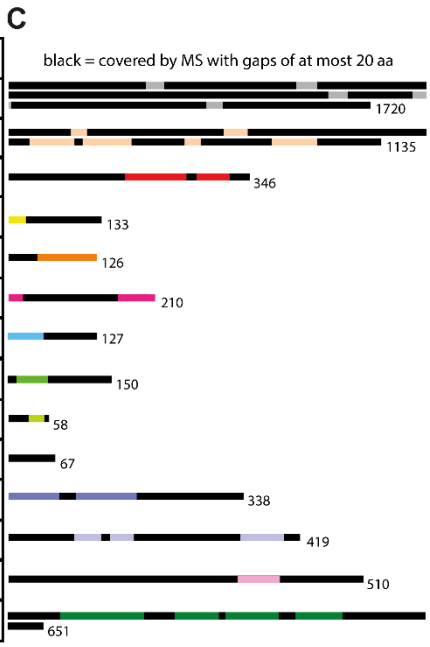


Figure 9 Mass spectrometry analysis of hPol I

A Coomassie-stained SDS-PAGE of human Pol I purification. **B** MS-results for the subunits and initiation factor RRN3 labeled in Panel (A) with sequence coverages of at least 25%. Number of identified peptides and score are indicated. **C** Schematic representation of sequence coverage, to scale from the N-terminus to the C-terminus. Black bars indicate covered sequences (with gaps <20 residues).

2.3 Evolutionary conservation of subunits and phylogenetic analysis of hPol I

Evolutionary conservation of Pol I subunits was first examined by structure-based alignments of human and yeast (*S. cerevisiae*) proteins. Figure 10 shows a schematic overview of subunits including domains as well as insertions and deletions larger than 10 aa compared between yeast and human. Larger divergence exists in clamp head and foot domain of RPA1, protrusion and clamp of RPA2, connector domain of RPA43, C-terminal region of RPA34 and RPABC2, whereas other subunits are mainly conserved in length and domain architecture.

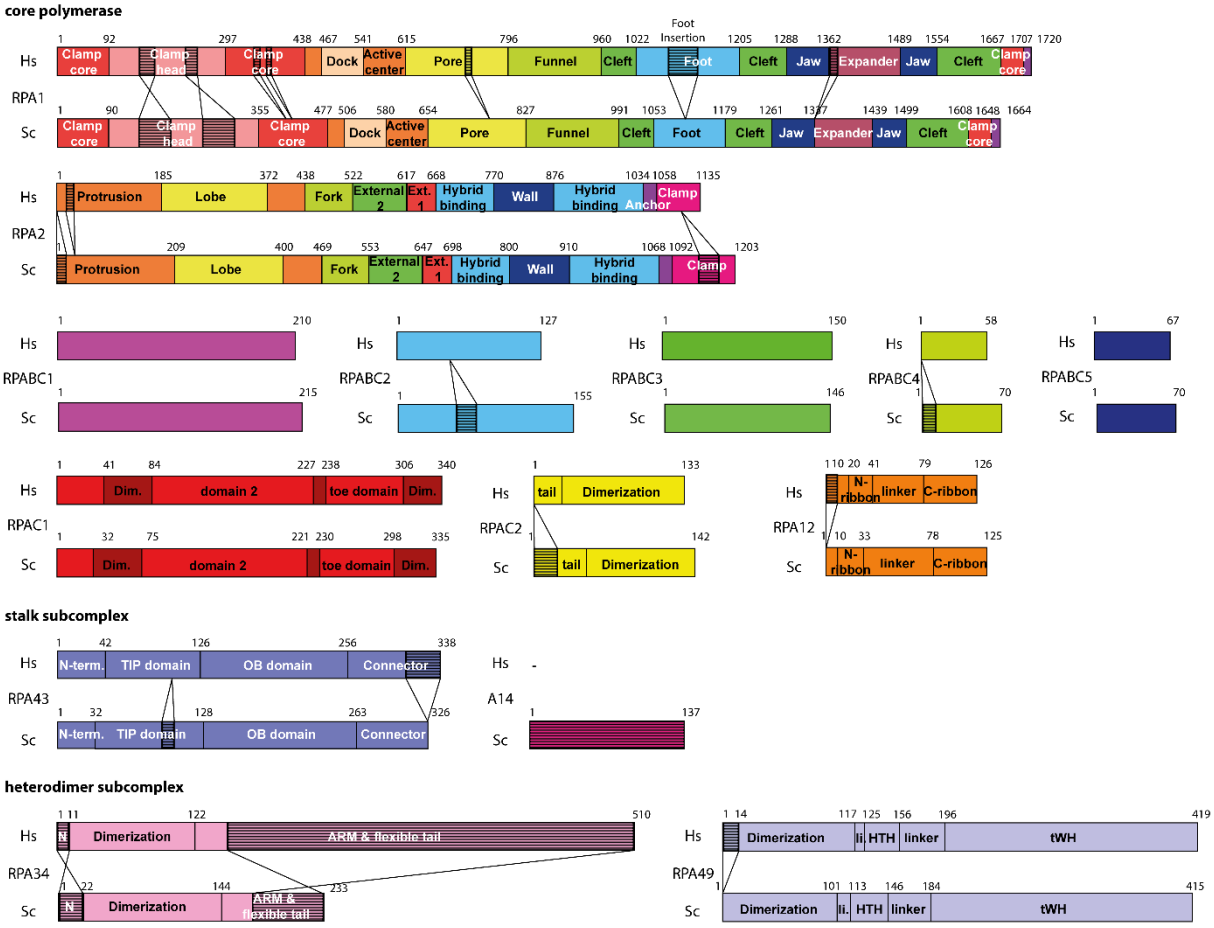


Figure 10 Human Pol I domain architecture in comparison to yeast Pol I

Schematic domain architecture of the hPol I subunits with labeled insertions and deletions in comparison to their Sc Pol I counterparts. Subdomains and insertions/deletions of 10 or more residues are indicated.

In yeast, Pol I is composed of 14 subunits, whereas in humans sequence-based searches on DNA and protein levels as well as our results from hPol I purification and mass spectrometry suggest a 13-subunit configuration[17]. Pol I subunit composition of different organisms throughout evolution was examined using phylogenetic analysis (see method section 6.14.1). For organisms included in the analysis, subunits RPA1, RPA34 and RPA43 (as representatives of Pol I core, heterodimer and stalk) were identified by sequence similarities. Based on the alignments of these sequences, a Pol I-specific conservation tree was generated removing bias which may originate from the influence of unrelated genes on global alignments in standard phylogenetic analysis. Additional to the three subunits, on which the organism selection was based, the occurrence of A14 was annotated (Figure 11A). The results showed a lack of subunit A14 in most organisms resulting in a single-subunit stalk. The predominant Pol I configuration was shown to be a 13-subunit complex. Only the *Saccharomycotina* in the clade of *Dikarya* harbored the 14th subunit A14 (44 of 513 included organisms). This observation also explained the paradigm why Pol I was believed to be a 14-subunit enzyme: Both investigated model systems *S. cerevisiae* and *S. pombe* belong to the *Saccharomycotina* group.

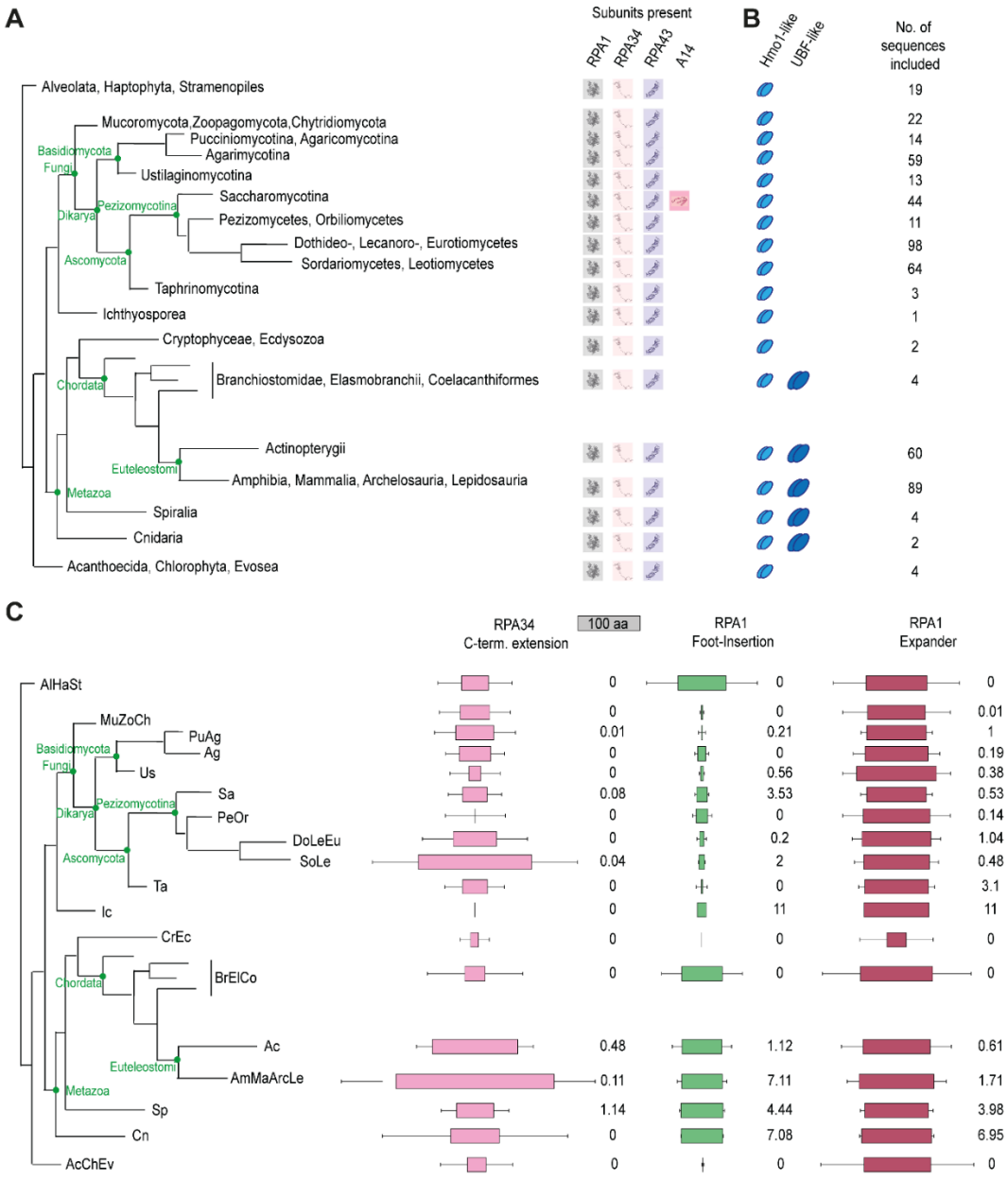


Figure 11 Phylogenetic analysis of Pol I

A Schematic phylogenetic tree calculated based on sequence homology of the three Pol I subunits RPA1 (core), RPA34 (RPA49/34 heterodimer), and RPA43 (stalk subcomplex). The subunit A14 was found in all *Saccharomycotina* in the class of *Dikarya*. This included model organisms such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, explaining the current paradigm that Pol I comprises 14 subunits. **B** HMG box domains (IPR009071) were present in Hmo1 as also in UBF versions. As in both proteins, a minimum of two consecutive boxes are present, the analyzed entries were restricted on such specific architecture. UBF-like proteins show five or six consecutive HMG boxes. Thus, the distribution of UBF proteins was restricted on the availability of the HMG box 5 (IPR029215). Taxonomic distribution of both entries was analyzed. A minimum of two consecutive HMG boxes could be found thoroughly over all organisms within our phylogenetic tree. The more specific HMG box 5 (UBF versions) was found within the higher order group of *Metazoa* (*Cryptophyceae* placed within *Metazoa* but belongs originally to the division of *Cryptophyta*). *Ecdysozoa* seemed to be more divergent from the group of *Metazoa* as only 1.5% of the sequences were annotated to have an HMG box 5. **C** Blocks show the median length of the C-terminal extension of RPA34, the RPA1 foot insertion, and the Expander (DNA-mimicking loop) (100 residue referenced above). Conservation scores were calculated for each specific region in every class. Box reflects the median; error bars indicate SD (standard deviation); conservation scores shown next to the boxes of each investigated element are grouped into five categories: not conserved (0 - 3), weakly conserved (3 - 5), medium conserved (5 - 7), conserved (7 - 9), and strongly conserved (9 - 11).

As mentioned in the introduction (see introduction section 1.4), the main difference between yeast and human initiation factors is the occurrence of UBF in humans, which is not present in yeast[157]. The closest homologue to UBF is Hmo1 in *S. cerevisiae* containing two HMG-boxes[79] (compare Figure 3). For all organisms included in the phylogenetic analysis, the presence of Hmo1 (at least two consecutive HMG-boxes) and UBF (five or six consecutive HMG-boxes and containing a homologue to HMG-box 5 of human UBF) versions were annotated (Figure 11B). A minimum of two consecutive HMG boxes was found thoroughly over all organisms within the Pol I-specific phylogenetic tree. The more specific HMG-box 5 (UBF versions) was found within the kingdom of *Metazoa* (also referred to as *Animalia*). Only the group *Cryptophyceae* and *Ecdysozoa,* which were combined and placed together within the *Metazoa* in the Pol I-based phylogenetic tree, seemed to be more divergent. *Cryoptophyceae* mostly contained a UBF version, but originally belong to the division of *Cryptophyta*, whereas only 1.5 % of the *Ecdysozoa* (belonging to *Metazoa*) possessed a UBF version.
In the next step, we analyzed differences within Pol I subunits. The expander and foot domain of RPA1, as well as the C-terminal region of RPA34 were further examined concerning their length and conservation score within different groups of the phylogenetic tree (Figure 11C). The DNA-mimicking expander loop of RPA1 showed overall the same length in all organisms, but rather low sequence conservation scores. However, the well-defined length of this element could hint to a conserved functionality. The insertion within the foot domain of RPA1, first observed between yeast and human, was distinct in length and seemed to be metazoan-specific. Subunit RPA34 is enlarged to 55 kDa (510 aa) in humans compared to 27 kDa (233 aa) in yeast, predominantly because of an altered and extended C-terminal part. For the C-terminal region of RPA34 the analysis showed no similarities concerning length or amino acid sequence within one group and comparing them among each other. Overall, a larger C-terminal extension was present in higher organism classes, such as Mammalia and Amphibia, but showed no clear conservation in sequence, predicted secondary structure, or length (Figure 11C). Hence, the RPA34 CTR was suggested to be a highly divergent Pol I element.

2.4 DNA-binding activity of subunits RPA49/34

Due to the divergence in the CTR of RPA34, the heterodimer RPA49/34 of hPol I may have differences to its yeast counterpart. Functional studies have already shown the important role of the heterodimer for Pol I transcription in different organisms, its DNA-binding activity, and differentially stable association to the core polymerase[56,60,65,66,102,177]. Recombinantly expressed and purified human heterodimer RPA49/34 was tested for DNA-binding in comparison to its yeast counterpart. For this purpose, yeast and human full-length (FL) $RPA49^{FL}/34^{FL}$, as well as human heterodimer with truncated RPA34-CTR ($RPA49^{FL}/RPA34^{1-343}$) and only the CTR of RPA34 ($RPA34^{131-510}$) were expressed and *in vitro* purified to > 95 % homogeneity (Figure 12A, B, see method sections 6.8.1 and 6.8.2). Electrophoretic mobility shift assays (EMSAs, see method section 6.12.1) were performed with these heterodimer variants using a 40 bp dsDNA fragment with a random sequence (Figure 12C). As expected, full-length RPA49/34 bound DNA possessing a similarly affinity for yeast and human heterodimer. As the main DNA-binding interface of the heterodimer is the tWH of subunit RPA49[5,60], it was not surprising that human heterodimer with truncated C-terminal region of RPA34 ($RPA49^{FL}/RPA34^{1-343}$) bound DNA. Interestingly, the flexible

C-terminal part of human RPA34 (RPA34$^{131-510}$) was also capable of independent DNA-binding itself (Figure 12C).

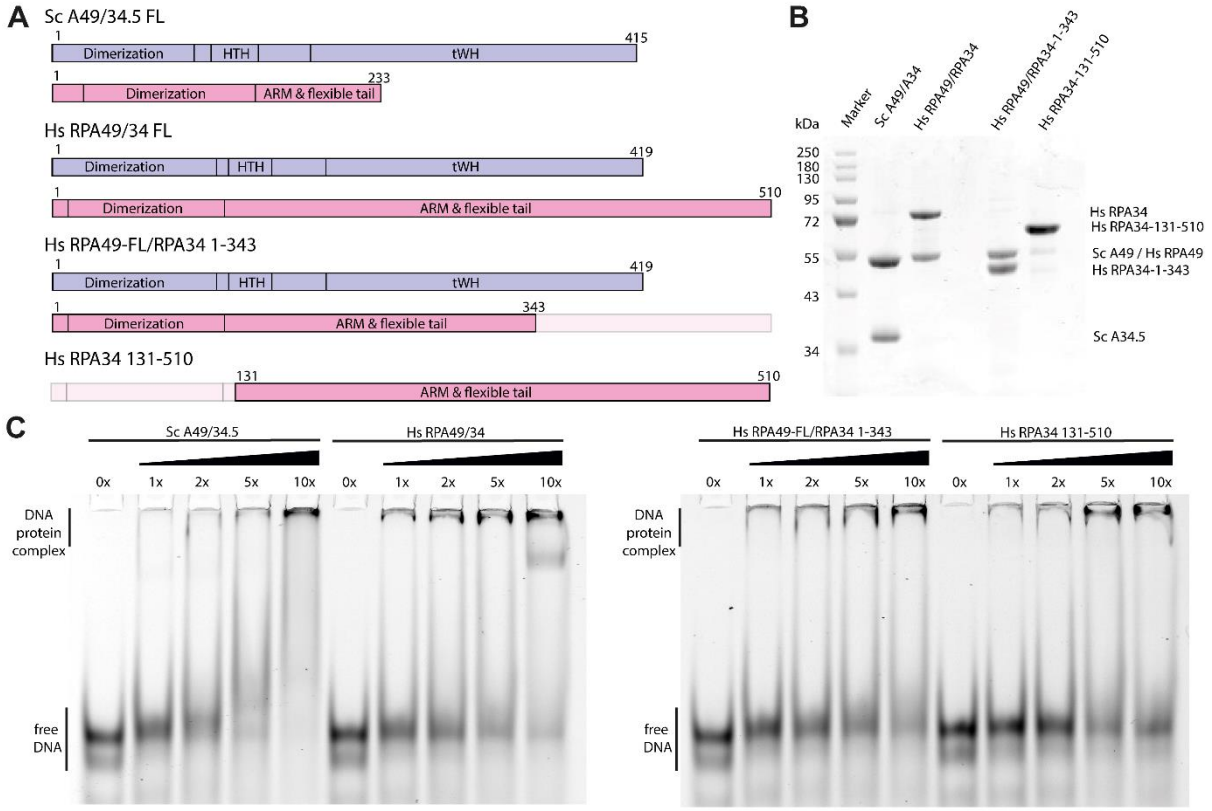


Figure 12 DNA-binding activity of the heterodimer RPA49/34

A Schematic representation of purified heterodimer variants. **B** Coomassie-stained SDS-PAGE of purified heterodimer variants. **C** Fluorescently labeled DNA was separated via native PAGE and detected DNA showed results of EMSAs. Bands of free and bound DNA visualized binding of purified heterodimer variants to a random 40 bp dsDNA. For all four purified heterodimer variants RPA49^{FL}/RPA34^{FL} of yeast and human, hs RPA49^{FL}/RPA34$^{1-343}$ and hs RPA34$^{131-510}$ DNA binding was shown.

2.5 Purified hPol I was functionally active in *in vitro* transcription

2.5.1 hPol I was active in elongation and cleavage

Purified yeast Pol I was extensively studied in numerous *in vitro* transcription assays examining subunit (sub-)domain functionalities in different transcription cycle states, elongation speed, or the interplay with transcription factors[51,54,61,64,65,98,130,186]. Equipped with the HeLa-POLR1A cell line and *in vitro* purified hPol I, similar *in vitro* functional analysis of the human enzyme and the comparison to the yeast model system could be accomplished. Firstly, hPol I activity was investigated in transcription elongation and cleavage *in vitro* assays (see method section 6.12.2). For this purpose, a minimal transcription scaffold composed of a short double-stranded (ds) DNA with a single-stranded (ss) overhang on the template strand was used. A fluorescently labeled RNA primer was annealed to the ssDNA overhang (Figure 13A). Using this scaffold, hPol I elongated the RNA primer upon addition of NTPs,

whereas it cleaved the RNA primer and backtracked without any NTPs in the reaction. Cleavage activity of Pol I relies on the activity of subunit RPA12, which harbors homology to the Pol II transcription factor TFIIS[52]. *In vitro* functional activity of purified hPol I was demonstrated including elongation in presence of NTPs and cleavage of the RNA primer in the absent of NTPs (Figure 13A). Specifically, elongation could be observed as an extension of the RNA primer (Figure 13A, lane 13-16) compared to the initial RNA primer (Figure 13A, lane 1) and cleavage as an RNA band shorter than the unprocessed RNA primer (Figure 13A, lane 12).

2.5.2 hPol I showed a reduced proof-reading activity compared to yeast Pol I

Phylogenetic analysis and subunit alignment already revealed differences between Pol I of different species including yeast and human leading to the question if *in vitro* functionality might differ (see result section 2.3). Thus, hPol I activity was compared to its yeast counterparts from the organisms *S. cerevisiae* (Sc) and *S. pombe* (Sp). The Pol I enzymes of the three different organisms were analyzed using the same transcription assay as above. NTPs were titrated into the reaction and the different RNA products were analyzed on a sequencing gel. In case of 'no NTPs', hPol I induced a cleavage band of -1 nt (Figure 13A, lane 12), whereas both yeast Pol I additionally showed a -3 nt band (Figure 13A, lane 2 and 7). This observation hinted to a reduced backtracking and/or cleavage activity of hPol I compared to its yeast counterpart. Upon NTP addition, transcription elongation was possible until a specific stalling site was reached dependent on the availability of specific NTPs. Sc and Sp Pol I stopped very specifically at the expected sites (Figure 13A, lane 3-5 and 8-10), whereas a remarkably larger portion of hPol I enzymes ran over the specific stalling site and hence had incorporated incorrectly paired nucleotides (Figure 13A, lane 13-15). Specifically, comparing the situation after the addition of the first nucleotide (ATP), the polymerases should have incorporated two nucleotides before having been stalled. For Sc and Sp Pol I the dominant band of extended RNA is +2 nt, whereas +2 nt and +3 nt bands were both main extension products of hPol I (Figure 13A, lane 3, 8, and 13). An explanation for these results may be either a reduced backtracking and/or cleavage activity in case of mis-incorporated nucleotides, an enhanced incorporation of incorrect nucleotides (promiscuity) or a combination of both in hPol I compared to its yeast counterpart. Taken together, this will be called a reduced proof-reading activity. Different minimal transcription scaffolds were tested to further examine these differences in activity (Figure 13B-F). An artificial transcription bubble, a shorter RNA primer, an altered distance between the end of the RNA primer and the beginning of the non-template strand, as well as a completely different sequence as minimal scaffold and bubble scaffold were used as transcription scaffolds. These variations resulted in the same overall effect of apparently reduced backtracking or cleavage and reduced proof-reading activity of hPol I compared to its yeast counterparts persisting with all tested scaffolds, although the effect seemed to vary in extent. For bubble scaffolds including a complete non-template strand the extent appeared to be less pronounced, but this potential trend needs to be verified by further experiments.

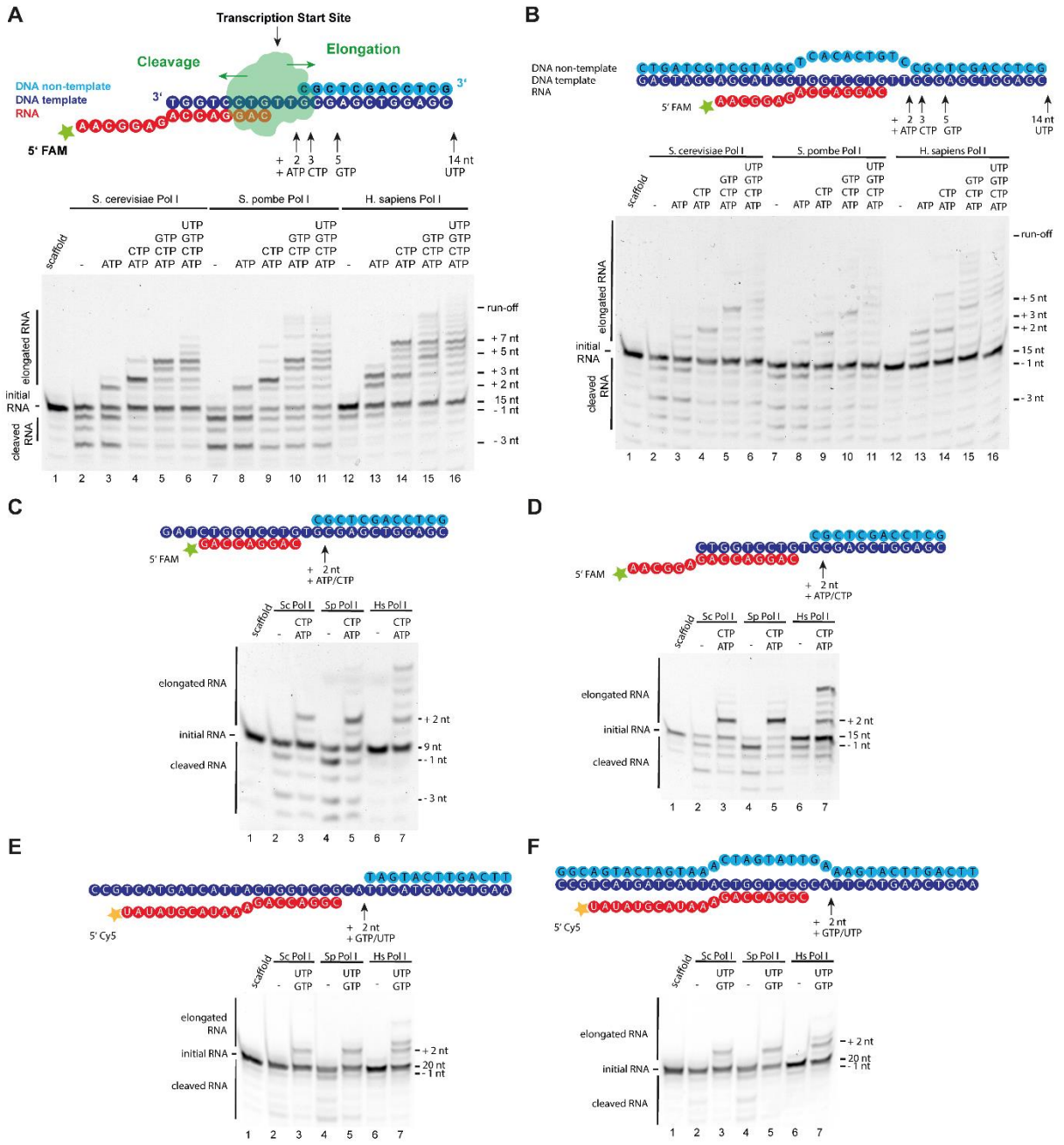


Figure 13 Functional activity of human Pol I compared to its yeast counterparts

A-F Schematic representation of the assay and scaffold sequence used to determine hPol I activity *in vitro*. Underneath the schematic representation, the sequencing gel (20 % Urea-PAGE) of the assay is shown. **A** RNA bands (detected via their fluorescent label) showed cleavage activity of hPol I by an additional RNA band at the -1 nt position, whereas Sc and Sp Pol I showed RNA cleavage bands at -1 nt and -3 nt position. Elongation of RNA by all polymerases was shown by longer RNA bands (+2/3/5/7 nt) compared to the initial RNA of 15 nt. **B** A bubble transcription scaffold (schematically shown on top) was used in the identical assay as in (A) and verified the results. **C-F** Different transcription scaffolds with a shorter RNA primer, an altered distance between the end of the RNA primer and the beginning of the non-template strand, as well as a completely different sequence for the minimal and the bubble transcription scaffold showed cleavage and elongation bands for the three different Pol I without and with NTPs, respectively.

2.5.3 Human subunits RPA49/34 and transcription factor RRN3 were not responsible for the reduced proof-reading activity

As the purifications of yeast and human Pol I showed some differences (see result section 2.2), it was examined whether these differences could be the reason for the differences in *in vitro* transcription (see result section 2.5.2). Thus, the influence of transcription initiation factor RRN3 and the heterodimer RPA49/34 were analyzed, as these factors were two main differences in yeast and human Pol I purifications.

Purified hPol I contained RRN3 to a higher extent than Sc and Sp Pol I. Purified Sc Rrn3 (Figure 14B) was added with increasing amounts to Sc Pol I whereupon elongation and cleavage activity was compared to the human counterpart using the same assay as described above (Figure 14A, C). Up to 5x molar excess of Rrn3 did not change Sc Pol I activity at all, whereas hPol I containing only sub-stochiometric amounts of RRN3 behaved differently, specifically showing a reduced backtracking and cleavage activity as well as an enhanced read-through activity at the stalling site. These experiments suggested that the transcription initiation factor RRN3 and its different occupancy was not the reason for the differences in activity between yeast and human Pol I.

As subunit RPA34 of the heterodimer is one of the most prominent differences between yeast and human (see Figure 10 and appendix section 7.1) and A49/34.5 mutations/deletions were shown to influence Sc Pol I activity[51], this sub-complex was investigated. First, human heterodimer RPA49/34 was titrated to hPol I using the same assay as described above to exclude the possibility that sub-stochiometric amounts of heterodimer in the hPol I purifications might result in the differences between yeast and human transcription activity (Figure 14D). As an increasing amount of added heterodimer did not alter neither the cleavage nor the elongation band pattern, the potential sub-stoichiometry of heterodimer was excluded as potential reason for the differences between yeast and human. The different source organisms of the heterodimer as potential difference was tested by adding yeast or human heterodimer RPA49/34 to hPol I ΔRPA49/34. This 11-subunit hPol I barely showed any transcriptional elongation or cleavage activity (Figure 14E, lane 1-2). Upon addition of Sc or Hs RPA49/34 the cleavage and elongation pattern of hPol I was restored (cleavage band of -1 nt and read-through of the expected stalling site at +3 nt) (Figure 14E, lane 3-6). These experiments indicated that the heterodimer indeed had an influence on transcriptional activity, as its absence led to strongly reduced overall transcriptional activity, however its source organism was not the reason for an altered proof-reading activity between yeast and human.

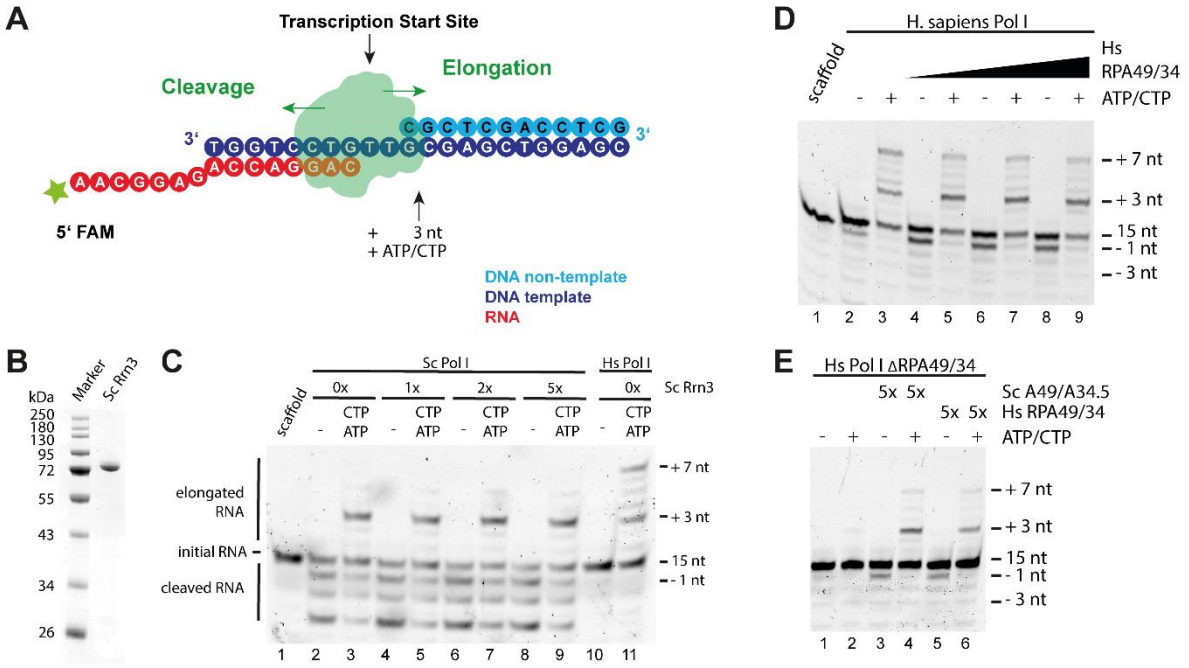


Figure 14 Influence of RRN3 and cross-species heterodimer on functional activity

A Schematic representation of the used transcription scaffold **B** Coomassie-stained SDS-PAGE of purified *S. cerevisiae* Rrn3. **C** RNA bands (detected via their fluorescent label) on sequencing gels (20 % Urea-PAGE) showed that the addition of yeast initiation factor Rrn3 did not impact elongation and cleavage functionality of Sc Pol I. For comparison, hPol I was used without Rrn3 addition. **D** Addition of recombinant human RPA49/34 did not abolish proof-reading deficiency nor induced deeper cleavage of human Pol I. **E** Human Pol I lacking the RPA49/34 heterodimer showed reduced cleavage and elongation activity. Complementation with recombinant human RPA49/34 or yeast A49/34.5 recovered elongation and cleavage activity but did not significantly change the pattern observed for complete hPol I (see Figure 14D).

2.6 Structure determination of hPol I

Yeast Pol I structures of different functional states are available, resolved by crystallography or cryo-EM[36,42,43,50,55,61,133], but structural information of the human enzyme was missing. Overall sequence and structure-based alignments as well as homology models suggested a global conservation of Pol I among organisms. We now aimed to investigate if and how species-specific differences may influence the enzyme and if structural divergence between yeast and human Pol I may explain functional differences. The *in vitro* purified hPol I made structural investigation possible. For structural analysis, cryo-EM was the method of choice due to the possibility to use low sample amounts with low concentrations and the availability of good molecular models from *S. cerevisiae*[42,43,187].

2.6.1 Structure of apo hPol I

Firstly, negatively stained sample of hPol I (see method sections 6.11.2 and 6.11.6) was examined. With this method, homogeneity and overall shape of hPol I was investigated. From 76 micrographs collected at a JEM-2100-F (JEOL) operated at 200 keV and equipped with a F416 CMOS-detector (TVIPS), about 46 000 particles were picked. After rough 2D classifications, 3D classification was performed (Figure 15). The resulting 3D reconstructions showed that a large portion of particles was intact and that in

some classes, densities from clamp and stalk regions were missing. This might indicate that hPol I possessed an enhanced flexibility of the clamp/stalk region compared to its yeast counterpart, but could have also resulted from damaged particles. Mass spectrometry results had shown that complete 13-subunit hPol I enzymes were purified (Figure 9) and used for grid preparation. Particle damaging could have also occurred during grid preparation at the air-water interface. Enhanced flexibility of the clamp and stalk regions could have originated from conformational heterogeneity or functional flexibility, which is an important feature for proper Pol I activity[5]. Besides this potential flexibility of the clamp/stalk region, the negative stain EM density of hPol I was overall similar to its yeast counterpart in architecture. Heterodimer RPA34/49, which is (partly) dissociated in some yeast Pol I reconstructions[44,55], seemed to be quite stably associated in the human complex as most polymerases showed density for the heterodimer sub-complex (Figure 15C).

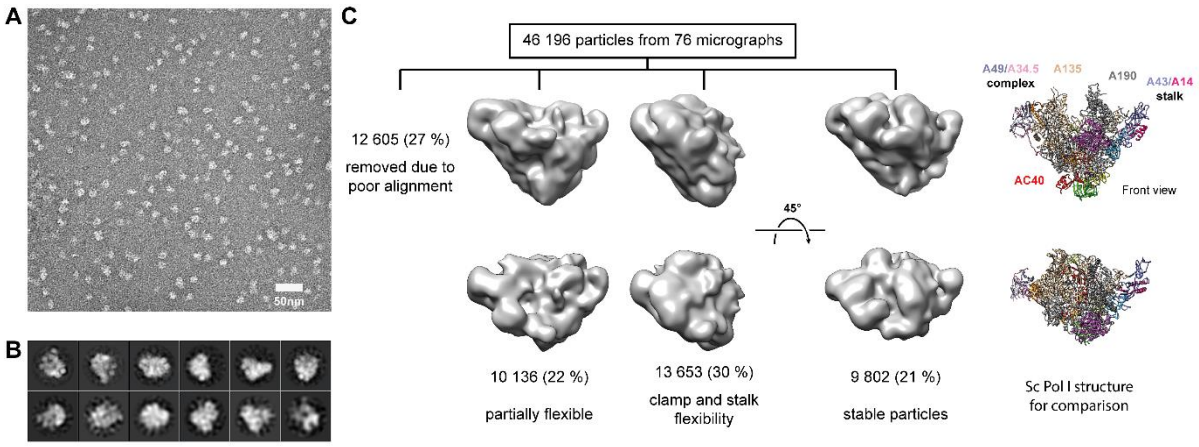


Figure 15 Human Pol I negative stain EM

A Exemplary negative stain EM micrograph of hPol I eluted from anti-GFP-nanobody beads. **B** Exemplary 2D classes of picked particles. **C** Processing of negative stain EM data set by 3D classification showed flexible particles and ~ 21% of intact, stable particles using EM of negatively stained sample. Two orientations of each 3D class are shown. The model of *Saccharomyces cerevisiae* Pol I is shown for comparison (right panel).

High-resolution structure determination was performed using cryo-EM (see method sections 6.11.5 and 6.11.7). For the collection of the first datasets, hPol I after anion-exchange chromatography was used in order to avoid 3C protease contamination and thus improving noise-to-signal ratio. Carbon-supported grids with monomeric hPol I were prepared and a dataset of 6 185 micrographs was collected on a Titan Krios with a K3 camera. After 2D and 3D classification 476 000 particles were refined. Unfortunately, a strong orientational bias already apparent in 2D classes prevented any meaningful 3D reconstruction of hPol I on carbon-supported grids due to a Fourier completeness of below 95 %. Merging these particles with a second dataset collected at a 30° tilt angle did improve the overall density a bit, but calculated Fourier completeness did not improve to gain a meaningful density. Different alterations in sample preparation were tried to overcome the orientational bias and several datasets with hPol I crosslinked with glutaraldehyde, containing different detergents in the buffer or with a tilt angle were collected. None of these variations resulted in appropriate improvements. The breakthrough was achieved using complete hPol I from the GFP-pulldown and graphene oxide (GO)-coated grids after extensive screening of different grid preparation methods[188]. Using these conditions

orientational distribution finally improved so far that a 3D reconstruction of an EM-density was possible. A total of 9 709 micrographs of GO-coated grids with non-crosslinked hPol I were collected on a CryoARM200 electron microscope (JEOL) equipped with a K2 direct electron detector (Gatan) with a dose of 40 $e^-/Å^2$ and a pixel size of 0.968 Å. After motion correction and CTF estimation, particles were picked and binned using WARP[189]. These particles were imported to RELION 4.0[190] for further processing. After 2D classification and sequential 3D classification a stable set of about 108 000 particles was selected and refined to an overall resolution of 4.09 Å after several rounds of 3D refinements, CTF refinements and map sharpening (Table 2, Figure 16).

Table 2 Cryo-EM data collection and refinement statistics

	Human Pol I
PDB	8A43
EMDB	15135
Data collection and processing	
Magnification	50.000
Voltage (kV)	200
Electron exposure ($e^-/Å^2$)	40
Defocus range (μm)	-1.2 - -2.7
Pixel size (Å)	0.968 (binned to 1.5085)
Symmetry imposed	C1
Initial particles images (no.)	145,554
Final particle images (no.)	108,012
Map resolution (Å)	4.09
FSC threshold	0.143
Refinement	
Initial model used (PDB code)	5M3M
Model resolution (Å)	3.5-4.1 Å
FSC threshold	0.143
Model composition	
Non-hydrogen atoms	31,110
Protein residues	3,914
Nucleotides	-
Ligands	-
B factors ($Å^2$)	
Protein	212.67
Nucleotides	-
Ligand	-
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.005
Validation	
MolProbity score	2.54
Clashscore	23.93
Poor rotamers (%)	0.41
Ramachandran plot	
Favored (%)	84.64
Allowed (%)	15.16
Disallowed (%)	0.21

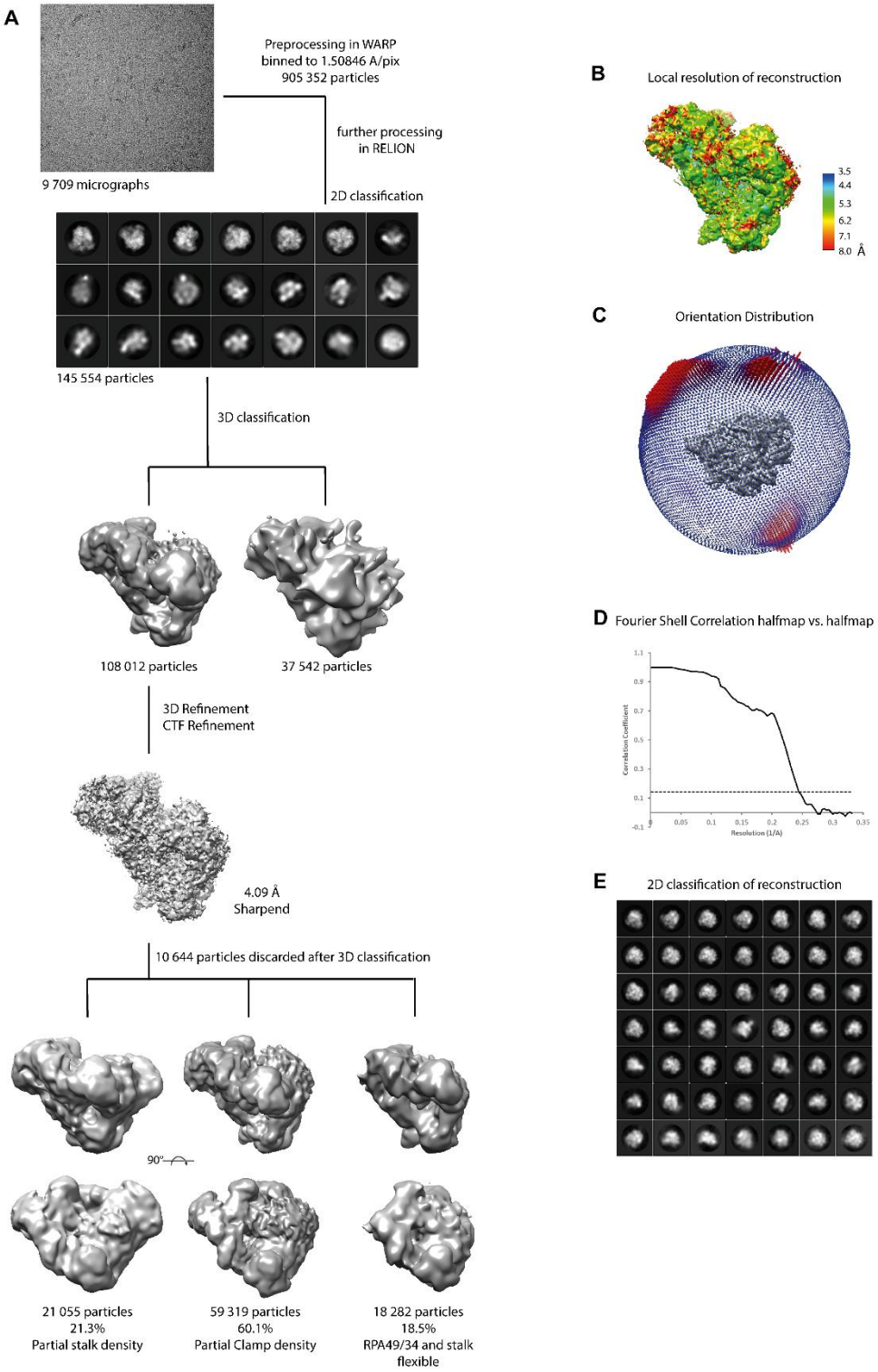


Figure 16 Processing tree of hPol I cryo-EM reconstruction

A Starting with 9 709 micrographs, preprocessing and particle picking was performed in WARP[189] and particles imported in RELION 4.0[190] for further processing. 2D and 3D classification led to a set of 108 012 particles, which were further processed by several rounds of CTF refinement and post-processing resulting in a cryo-EM density of 4.09 Å. 3D classification of this final reconstructions indicated that only 21.3 % of particles showed partial stalk density. **B** Local resolution of the cryo-EM density showed higher local resolution of the core polymerase and lower resolution on the surface. **C** Angular distribution plot of particles contributing to the final reconstruction. **D** Fourier Shell Correlation (FSC) of final reconstruction using the gold standard of FSC threshold of 0.143. **E** 2D classification without sampling for final reconstruction.

Similar to the results from the EM density of negative stained sample, clamp core, clamp head, and stalk regions remained flexible. Further 3D classification revealed that the lack of density in this Pol I part was even enhanced in cryo-EM reconstructions as rather few particles showed partial stalk density in the best case (Figure 16, Figure 17). Connected EM density was visible for the polymerase core as well as for the dimerization domain of the sub-complex RPA49/34 except for the flexible clamp core and head. The single-subunit stalk sub-complex (RPA43), the C-terminal part of RPA12 and the C-terminal part of RPA34 showed a high extent of flexibility and were not visible. On the downstream site of the polymerase additional density was found which was located on top of RPABC5 and close to the RPA1 foot domain. For model building, the common (RPABC1-5) and Pol III shared (RPAC1/2) subunits were transferred from a hPol III model[191]. Homology models using the structural information from their *S. cerevisiae* counterparts[42] were calculated, using the MODELLER software package[192], for Pol I specific subunits RPA1, RPA2, RPA43, RPA49, RPA34 and RPA12. The computationally calculated model was fitted unambiguously into the cryo-EM density of hPol I and afterwards fitted in detail into the density manually. Further refinement was performed using Phenix[193] and inspected manually. The extra density on the downstream edge of the polymerase could be assigned to the human-specific insertion in the foot domain. Model building of this domain was further assisted by its AlphaFold[194] prediction and domain prediction using HHPRED[195]. The latter unraveled a homology to an HMG-box domain with the closest relation to the structure of HMG-box 5 of human transcription factor UBF[14]. Without doubt, the model for the foot insertion could be built into the density and an HMG-box fold was clearly assigned to the now named 'dock II' domain. Hence, a cryo-EM reconstruction of monomeric apo hPol I was obtained including RPABC1-5, the RPAC1/2 dimer, the N-terminal lobe-binding domain of RPA12, the dimerization domain of RPA34/49, most parts of RPA2 except of the C-terminal clamp and anchor domains (aa 1010-1134), as well as jaw, funnel, foot, and most of the cleft domains of RPA1 (aa 630-1661 excluding some loops). The C-terminal domain of RPA12 responsible for the intrinsic cleavage activity as well as the C-terminal extension of RPA34 stayed flexible, similar to many reconstructions from yeast Pol I[36,44,50]. Linker and tWH of RPA49 were flexible, as expected from yeast apo Pol I structures[36,55]. Very weak density was found for the stalk (RPA43) as well as for clamp and dock domains of RPA1 indicating an enhanced flexibility of the shelf module. Overall architecture of Pol I was consistent between yeast and human, but specific mapping of disease-causing mutations in humans and their interpretation could be performed in more detail and adequately only with the structure of the human enzyme.

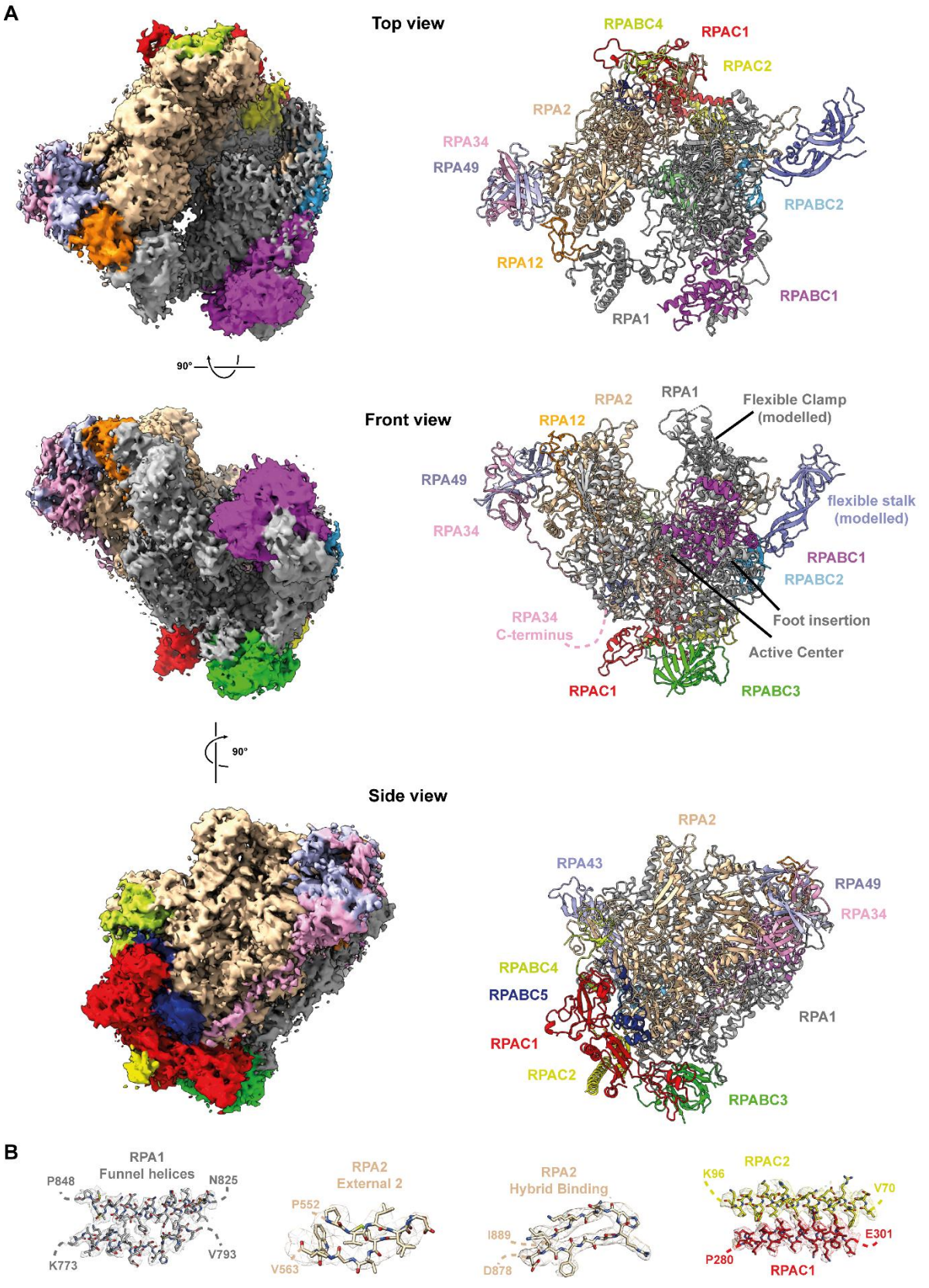


Figure 17 Cryo-EM reconstruction of monomeric hPol I

A Cryo-EM density of human Pol I showed flexibilities in the clamp/stalk region of RPA1 and RPA43. The structural model in cartoon representation is shown besides. **B** Enlarged view of RPA1 funnel helices, RPA2 External 2 and Hybrid Binding domains, and the RPAC1/2 assembly overlaid with sharpened cryo-EM density.

2.6.2 Mapping of disease-causing mutations on hPol I structure

A number of mutations in the sequence of hPol I subunits have been identified to cause a variety of human diseases and could now be mapped on the human structure (Figure 18A, Table 3). Acrofacial Dysostosis (Cincinnati type)[196,197], Treacher-Collins syndrome (TCS)[198–201], Hypomyelinating Leukodystrophy (HL)[200,202], and a juvenile neurodegenerative phenotype akin to the HL-phenotype[203] are caused by mutations within different hPol I subunits.

Acrofacial Dysostosis, Cincinnati type, leads to craniofacial abnormalities during development and is caused by mutation E593Q or V1299F in subunit RPA1[196,197]. Mutation E593Q is located in proximity to the catalytic center and may directly affect the nucleotide addition, which is in line with a reported transcription suppression of this mutant due to its enhanced rDNA binding stability[196] (Figure 18C). Recently, E593Q-hPol I was shown to bind UBF-bound rDNA and thereby competitively inhibits wild-type hPol I[197]. In contrast, V1299F is situated on the interface of RPA1 with RPA12 and may destabilize the association of subunit RPA12 with the hPol I core (Figure 18D).

Treacher-Collins syndrome is another craniofacial developmental disease caused by various mutations in the genes *TCOF1* (encoding for the Pol I regulating protein Treacle), *POLR1B* (subunit RPA2), *POLR1C* (subunit RPAC1), or *POLR1D* (subunit RPAC2). Serine 682 of RPA2 directly contacts the bridge helix (likely H967 of RPA1), indicating that mutation S682R may affect the bridge helix resulting in partially hindered translocation (Figure 18F). In contrast, R1003 of subunit RPA2 is situated in the DNA/RNA binding cleft, takes part in intra-subunit contacts, and may be required to stabilize folding of the hybrid-binding domain within RPA2 (Figure 18G). Hence R1003C and R1003S[201] may lead to a destabilization of RPA2 and thus the active center. Many TCS-associated mutations of RPAC2 (E47K, T50I, L51R, G52E, L55V, R56C, L82S, G99S) structurally cluster together at an inter- and intra-subunit interaction site of RPAC1 and RPAC2[198–200] (Figure 18H). Structural alignment with the human Pol III structure revealed a similar fold and suggested destabilizing effects of these mutations, similar to R279Q/W of subunit RPAC1. Interestingly, R279Q only affects Pol I import, but neither Pol III assembly nor its import[200]. Nevertheless, it can not be excluded that some TCS-mutations effect both Pol I and Pol III transcription systems.

Similar to TCS, Hypomyelinating Leukodystrophy is a neurodegenerative disease that cannot be classified as a Pol I- or Pol III-associated disease *per se*. HL mutations are found in subunit RPAC1 which is shared between both polymerases or in Pol III subunits RPC1 and RPC2[200,202]. Comparing the structures of hPol I and hPol III showed that mutations of the RPAC1 N-terminus (T26I, T27A, P30S, N32I) were likely to have a Pol III-specific effect as this N-terminal region appeared flexible in hPol I, but mediated interactions to the polymerase core in Pol III (Figure 18I). N32I and the nearby mutation N74S have been demonstrated to affect Pol III assembly, nuclear import, and chromatin association but does not impair Pol I biogenesis nor nuclear import[200]. Additionally, RPAC1 mutations I105F, H108Y, and R109H were found to impair RPC2 interaction in hPol III but not RPA2 in Pol I, again suggesting Pol III-specificity (Figure 18J). RPAC1 mutations M65V, V94A, A117P, G132D, C146R, R191Q, I262T, T313M, and E324K are involved in the formation of intra-subunit contacts, likely affecting RPAC1 folding itself, and thereby might have a destabilizing effect on both enzymes (Figure 18E).

Finally, the mutation S934L in RPA1 is associated with a juvenile neurodegenerative phenotype akin to the HL-phenotype associated with Pol III disruption[203]. This mutation occurs in a small loop of RPA1 which forms contacts with RPA2 in the vicinity of the bridge helix N-terminus (Figure 18B). The destabilization of these contacts may generally disrupt and destabilize the Pol I core to some extent.

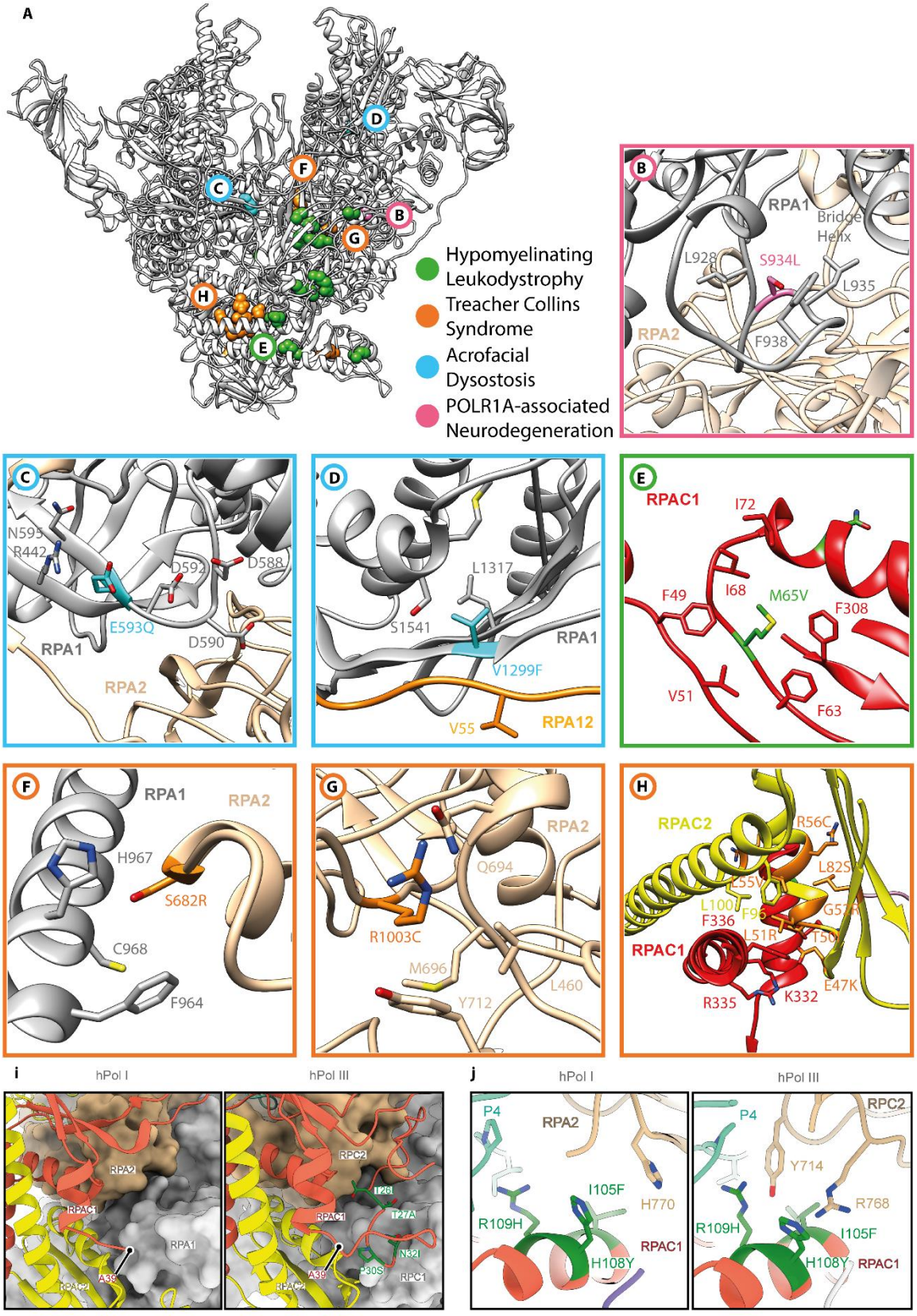


Figure 18 Mapping of disease-causing mutations with hPol I

A Structural model of hPol I (grey ribbon, back view) with location of disease-related mutation indicated (color code in panel). **B-H** Close-up views of the residues outlined in (A). **I** The RPAC1 N-terminal region is flexible in hPol I compared to hPol III. **J** Interface of RPAC1 residues 105-109 with the second largest subunit in hPol I and hPol III.

Table 3 Known mutations in Pol I related to human diseases

Mutation		potential impact on Pol I
Acrofacial dysostosis, Cincinnati type		
RPA1	V1299F[196]	might destabilize RPA1 fold intrinsically and RPA12 association
	E593Q[196]	might affect nucleotide addition cycle
Treacher Collins syndrome		
RPA2	S682R[201]	might destabilize the bridge helix
	R1003C/S[201]	might affect nucleotide association
RPAC1	R279Q/W[199,200]	might destabilize RPAC1 fold intrinsically
RPAC2	E47K[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	T50I[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	L51R[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	G52E[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	L55V[198]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	R56C[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	L82S[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
	G99S[199]	might destabilize RPAC2 fold intrinsically and RPAC1-RPAC2 interaction
Treacle protein mutations		mutations in a Pol I regulating protein
Hypomyelinating Leukodystrophy		
RPC1 mutations		affects Pol III transcription system
RPC2 mutations		affects Pol III transcription system
RPAC1	T26I[200,202]	might predominantly affect the Pol III system
	T27A[202]	might predominantly affect the Pol III system
	P30S[202]	might predominantly affect the Pol III system
	N32I[200,202]	might predominantly affect the Pol III system
	M65V[200,202]	might destabilize RPAC1 fold intrinsically
	N74S[200,202]	might predominantly affect the Pol III system
	V94A[200,202]	might destabilize RPAC1 fold intrinsically
	I105F[202]	might predominantly affect the Pol III system
	H108Y[202]	might predominantly affect the Pol III system
	R109H[200,202]	might predominantly affect the Pol III system
	A117P[202]	might destabilize RPAC1 fold intrinsically
	G132D[200,202]	might destabilize RPAC1 fold intrinsically
	C146R[200,202]	might destabilize RPAC1 fold intrinsically
	R191Q[200,202]	might destabilize RPAC1 fold intrinsically
	I262T[200,202]	might destabilize RPAC1 fold intrinsically
	T313M[202]	might destabilize RPAC1 fold intrinsically
	E324K[200,202]	might destabilize RPAC1 fold intrinsically
juvenile neurodegenerative phenotype akin to the HL-phenotype		
RPA1	S934L[203]	might destabilize contacts to RPA2

2.7 Human-specific dock II domain

The structure of hPol I uncovered the fold of the human-specific foot insertion, the dock II domain. This insertion of the largest subunit RPA1 folds to an HMG-box and is located on the downstream face of the polymerase (Figure 19). Dock II is in close proximity to RPABC1 and includes three inserted helices α27d-f compared to the yeast Pol I structure. Dock II is suggested to be *metazoan*-specific as indicated by the phylogenetic analysis (Figure 11C, see result section 2.3). In comparison to Pol I, Pol II

has also an enlarged foot domain in yeast[42,43]. Pol II foot domain is responsible for a transient interaction with the 'mediator' complex and involved in transcription regulation[204,205]. Having again an enlarged foot domain from yeast to human in Pol I could constitute to an additional regulatory feature in humans, therefore the dock II domain was further investigated.

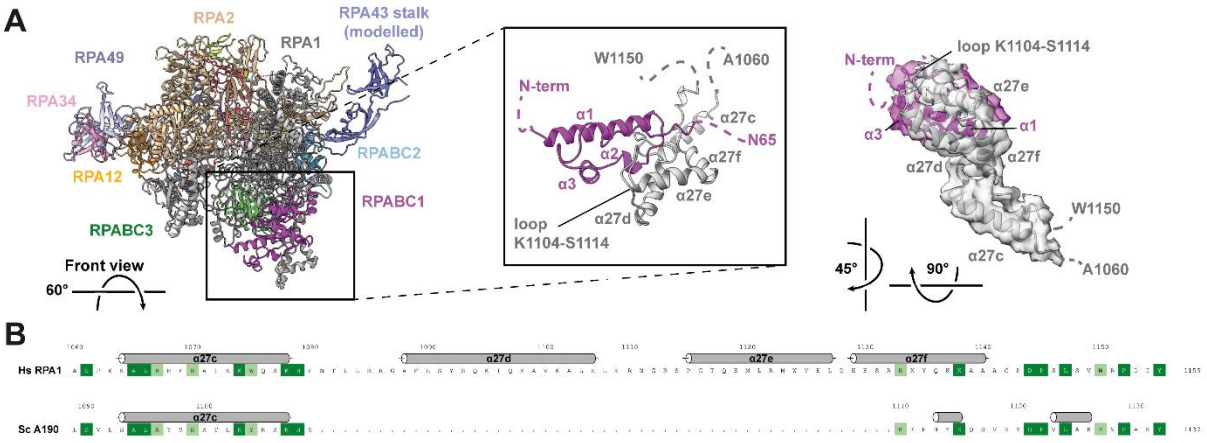


Figure 19 Human-specific 'dock II' domain

A Location of the structured insertion in RPA1 α27d-f on the downstream edge of subunit RPABC1 in human Pol I and enlarged view of the region. Overlaid experimental cryo-EM density for the helices α27c-f of subunit RPA1 (grey) and the N-terminal 65 residues of RPABC1 (purple) shown as transparent surface (right). **B** Structure-based sequence alignment of human and yeast Pol I foot insertions (for complete sequence compare section 7.1 structure-based alignments).

2.7.1 Dock II did not possess strong DNA-binding affinity

Many HMG-boxes function as DNA-binding domains[206], why this potential function of the dock II domain was first investigated. MBP-tagged full-length and minimal dock II domain as well as MBP-tag-only was *in vitro* purified (Figure 20A, see method section 6.8.5) and its affinity towards a random 40 bp long dsDNA fragment was examined using EMSAs (Figure 20B, see method section 6.12.1). Canonical HMG-boxes bind dsDNA in a sequence-specific or -unspecific manner to the minor groove with a preference to non-B-form DNA[206]. Overall DNA-binding of any dock II variant was absent or weak up to high concentrations of the dock II domain. Only the positive control, yeast core factor, bound DNA with a high affinity (Figure 20B). As for full-length dock II at high concentrations, a low amount of protein-DNA complex was detected, the electrostatic potential of the dock II domain as potential reason for the low affinity binding was calculated using ChimeraX[207] (Figure 20C). A large positively charged surface was found on the downstream site, which could explain some non-specific DNA-binding at high protein concentrations. Furthermore, the structure of the dock II domain was compared to the DNA-bound form of HMG-box 2 of the protein HMGB1 (PDB 6CIK)[208] (Figure 20D, E).

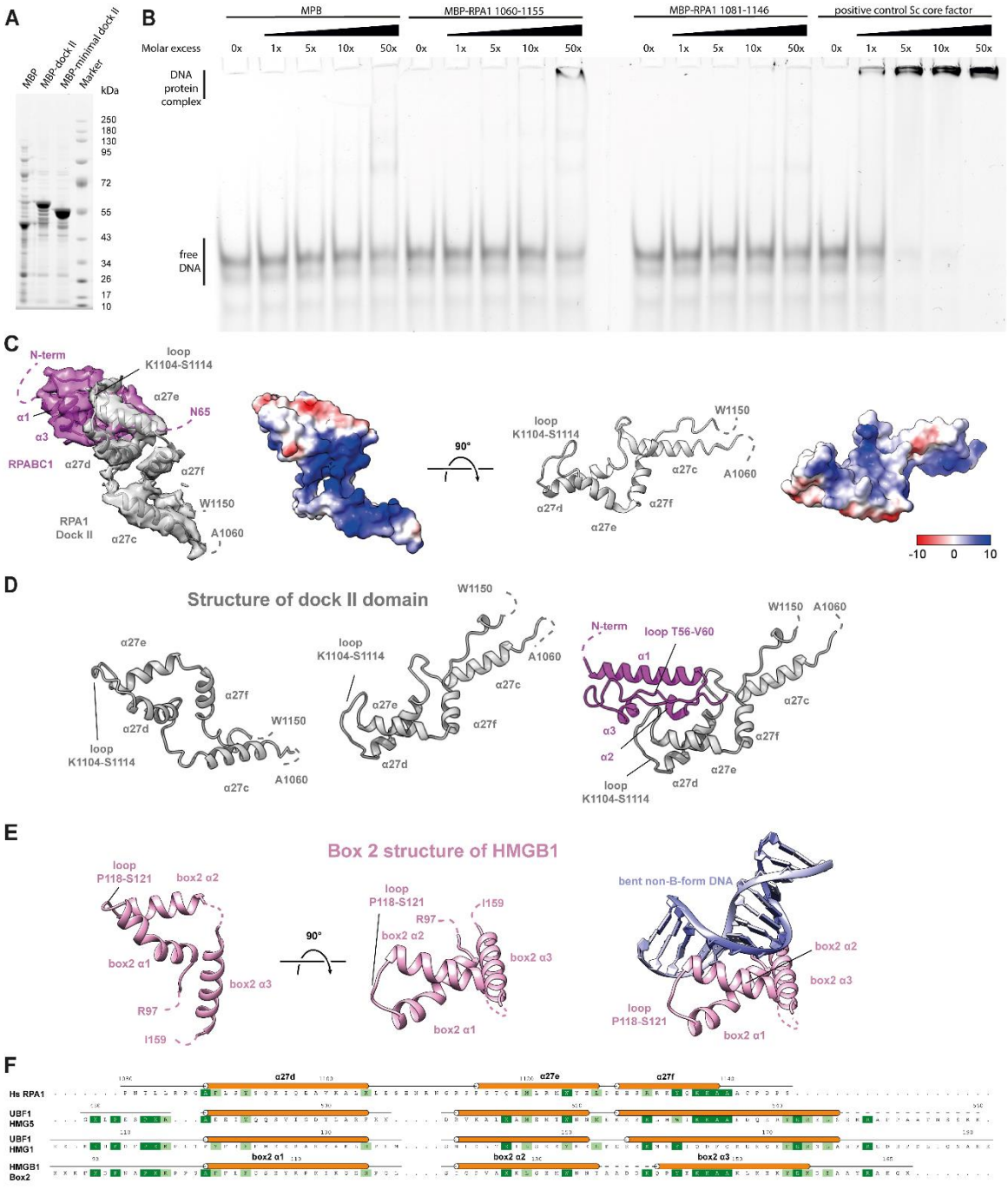


Figure 20 Dock II domain did not possess DNA-binding capability

A Coomassie-stained SDS-PAGE showed purified dock II-MBP fusion proteins. **B** DNA-protein complexes and free DNA were separated on 6 % native PAGE gels and the DNA was detected via its fluorescent label. The results of the EMSAs showed no or only low affinity of the different recombinant dock II versions to a 40 bp dsDNA fragment visualized by no or a weak shift of the free DNA band to higher molecular weight at high protein concentrations. Negative control: MBP-tag only, positive control: Sc core factor. **C** The electrostatic potential of the dock II domain calculated using ChimeraX[207] is shown in two different orientations. **D, E** Structure of the RPA1-foot insertion (D) compared with the canonical HMG box 2 of the human protein HMGB1 (PDB 6CIK)[208] (E) from two views. The DNA-binding surface of the canonical HMG box 2 is occluded by RPABC1 in hPol I. **F** Structure-based sequence alignment of the RPA1-HMG insertion with the canonical HMG box 2 of HMGB1, and the boxes 1 and 5 of the Pol I transcription factor UBF. In the RPA1-HMG box, the N-terminal region is divergent and the third helix is truncated, which are two features important for DNA-interaction. A loop insertion between the first two helices is part of the RPABC1 interface.

This comparison was related to the structure-based sequence alignment of dock II with other HMG-boxes (Figure 20F). The alignment showed that the third helix α27f was shortened compared to canonical DNA-binding HMG-boxes, while this helix is largely involved in the DNA-binding affinity of canonical HMG-boxes[209]. Furthermore, the structure-based sequence alignment of different HMG-boxes uncovered another difference: a motive N-terminal of the first helix is conserved among HMG-boxes but was missing in dock II. This motive builds the 'minor wing' of HMG-boxes together with the C-terminal region of the third helix[206], which was absent in dock II as mentioned before. The 'minor wing' of canonical HMG-boxes is usually DNA-binding via a DNA-binding cooperativity of its two above mentioned regions[209]. Additionally, the loop between helices α27d and α27e, which contributes to DNA-binding and sequence specificity in canonical HMG-boxes[209], was enlarged in the dock II HMG-box. The enlarged loop in RPA1 was instead interacting with the loop T56-V60 of subunit RPABC1. Going along with these observations, the DNA-interaction surface was occupied by RPABC1 in the case of the dock II HMG-box when the overall architecture and DNA-binding surface got transferred from HMG-box 2 of HMGB1 representing DNA-binding HMG-boxes to the dock II domain (Figure 20C, D). The resulting conclusion was that the dock II domain represents a degenerated HMG-box, forming a truncated 'major wing' and excluded a major DNA-binding function.

2.7.2 Dock II domain as a potential new protein-protein interaction platform

Besides the DNA-binding function of HMG-boxes, it is known that some HMG-boxes can act as protein-protein interaction platform[206]. As Topoisomerase 2a (Top2a) was found to be part of human Pol I transcription initiation complexes[150,210] on the one hand, and on the other hand HMGB1 can bind Top2a via its HMG-box and promote Top2a activity[211], the possibility of Top2a binding to the dock II domain was examined. We teamed up with Prof. Dr. Valerie Lamour and received *in vitro* purified Top2aΔCTD[212], lacking the unstructured C-terminal domain (Figure 21A). Immobilizing dock II versions on different beads and testing Top2a interaction using pulldowns (see method section 6.12.4) was not successful as Top2a bound unspecifically to any tested bead material in many different buffer systems (Figure 21B). As a bead-independent method, blue-native PAGE (see method section 6.12.5) was performed to investigate dock II - Top2a interaction (Figure 21C). For Top2a incubated with full-length dock II, a shifted Top2a band resembling dock II - Top2a complex was found, whereas no interaction was detected for minimal dock II or tag-only. To further verify band identities, mass spectrometry was performed (Figure 21D, see method section 6.14.2). Indeed, for the negative control Top2a band (band '§'), only Top2a was detected. The Top2a band of the full-length dock II - Top2a sample (band '#') contained a lot of Top2a as well as low amounts of MBP and dock II domain, whereas the shifted complex band (band '*') included all three parts in high amounts. These results suggested a transient low-affinity interaction between Top2a and human dock II domain, hence proposing this *metazoan*-specific domain as a potential protein-protein interaction platform.

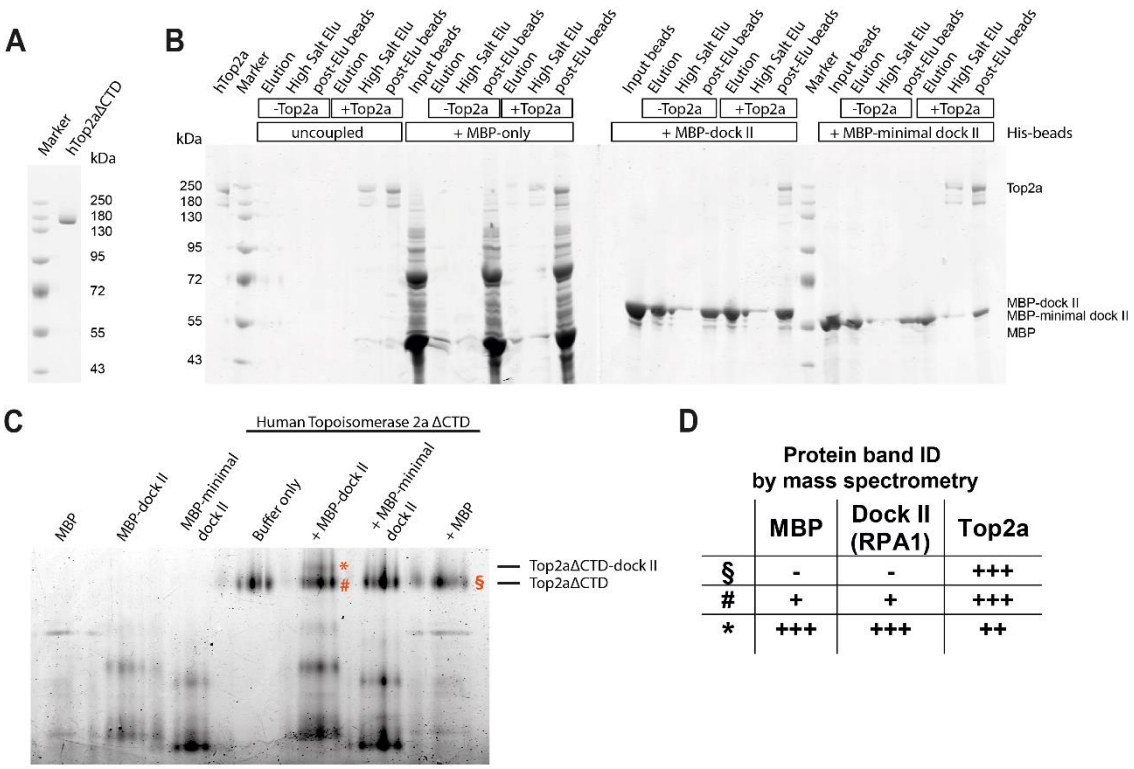


Figure 21 Dock II - Top2a interaction

A Coomassie-stained SDS-PAGE showed purified recombinant Top2a lacking the CTD. **B** Coomassie-stained SDS gel of an exemplary Top2a pulldown with uncoupled or MBP/MBP-dock II/MBP-minimal dock II - coupled beads showed that Top2a was found in all high salt elution and post-elution beads fractions independent from the bead sample (in case Top2a was added). Top2a seemed to be sticky to the bead material as elution with imidazole (elution fractions) and with high salt (high salt elution fractions) was not or only partially possible. Control samples without any Top2a addition showed that elution with imidazole (Elution fractions) was not complete hinting to the possibility of precipitation on the bead material. **C** Native PAGE analysis showed a shift in the main Top2aΔCTD band in the presence of recombinant full-length MBP-dock II indicating a low-affinity interaction. The shift was not observed in the minimal dock II-MBP fusion or the MBP-only lanes. **D** Mass spectrometry analysis for unshifted (#) and shifted (*) band of Top2a incubated with full-length MBP-dock II and for Top2a incubated with MBP-only (§) showed that Top2a was present in all bands. MBP and dock II was found in the shifted band and, in lower amounts, in the unshifted band of Top2a + MBP-dock II, but not in the Top2a band incubated with MBP-only.

To further test the possibility of direct binding of Top2a to hPol I via the dock II domain, *in situ* molecular modelling was performed (Figure 22). hPol I fragments (subunit RPA1) were docked to Top2a in its structurally resolved states I and II using HADDOCK[213], AutoDock Vina[214], the ZDOCK webserver[215], and PRISM webserver[216,217] (see method section 6.14.3). Interaction patterns were analyzed using the MAXIMOBY (CHEOPS) contact matrix algorithm and the VMD plugin yContact[218] (for details, see method section 6.14.3). The two most reliable docking results for RPA1 to both Top2a states, respectively, using the HADDOCK[213] software were further examined (Figure 22A). For all predicted interactions the dock II domain was a major interaction surface and no major clashes between hPol I and Top2a were detected. Interestingly, for all predicted structures, the hPol I downstream DNA path was freely accessible. Nevertheless, no common interaction surface was observed for Top2a, but interaction options for the catalytic domain (interaction option I and II in Figure 22A) and the ATPase domain (interaction option III and IV in Figure 22A) were predicted. *In situ* predictions for the interaction between dock II domain of hPol I and Top2a were generally in line with HADDOCK results of complete

RPA1 (Figure 22B). The different software tools predicted similar binding possibilities of dock II for both Top2a states within the catalytic and ATPase domain. Together the results of the blue native PAGE analysis and the *in situ* molecular modelling may suggest dock II as a direct interaction platform for Top2a.

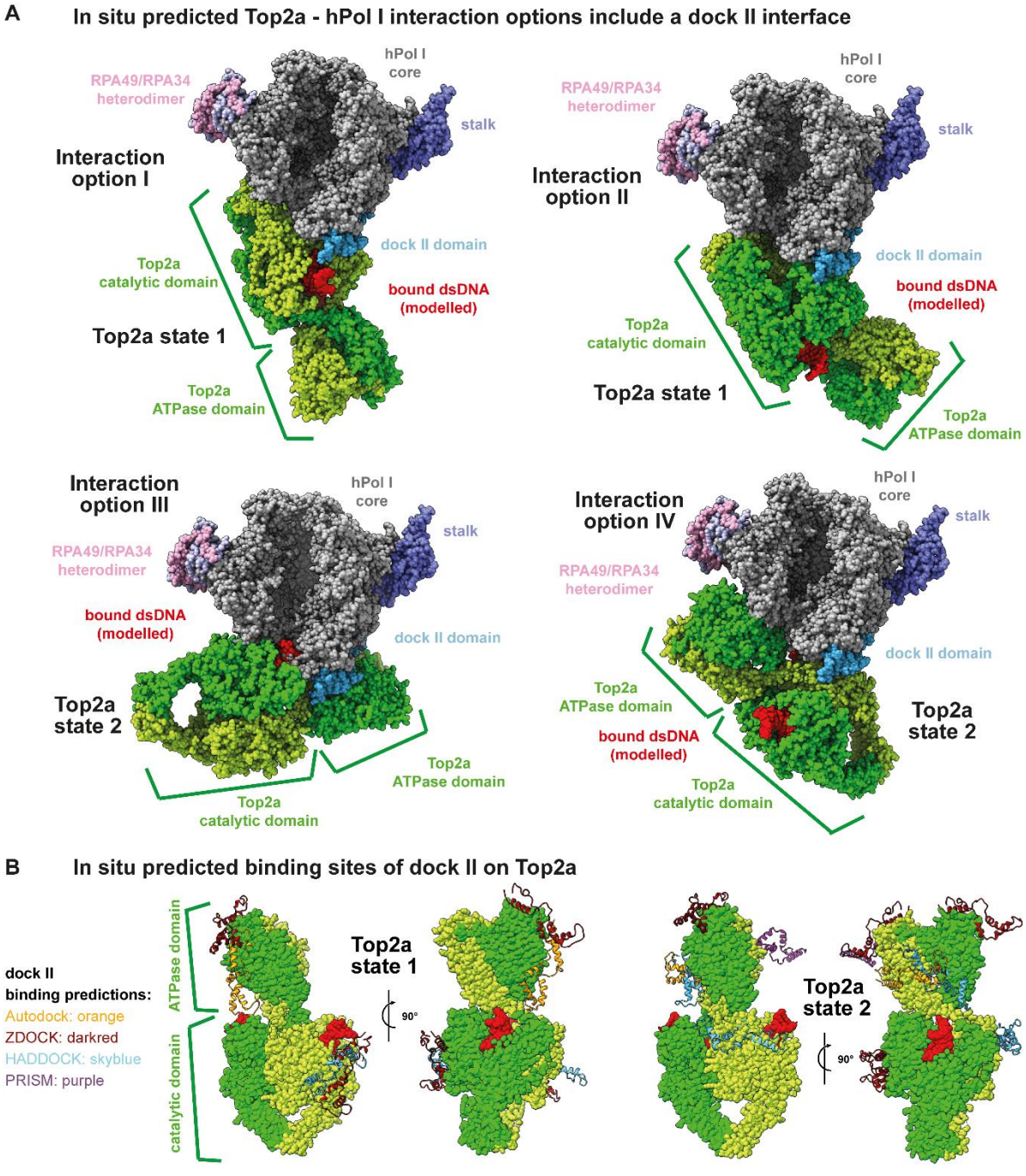


Figure 22 In situ modelling identifies the dock II domain as potential Top2a - hPol I interaction site

A Four different interaction options between Top2a in state 1 or 2 (subunits of homodimer shown in lemon and green, respectively; dsDNA in red) and hPol I subunit RPA1 predicted by HADDOCK are shown (hPol I core in grey; RPA43 stalk and RPA49/RPA34 heterodimer highlighted; dock II domain in sky blue; hPol I shown as front view tilted 60° towards top view). **B** *In situ* predicted binding sites of full-length dock II to Top2a (colors as in (A)) in state 1 and 2 (PDB 6ZY7 and 6ZY8, respectively) using HADDOCK (sky blue), Autodock (orange), ZDOCK (dark red), and PRISM (purple) are illustrated.

Having shown that a direct interaction between hPol I and Top2a might be possible *in vitro* and *in situ,* Top2a localization at the rDNA locus was examined as a pre-requisite for potential *in vivo* interaction. Thus, previously published Top2a ChIP-Seq data from mouse cells[219] were re-analyzed concerning rDNA localization. Additionally, Ubf, Pol I, Rrn3, Taf1b and TBP[19,71,220] were mapped to the rDNA locus (Figure 23A). SL1 components Taf1b and TBP were clearly enriched at the gene and spacer promoter as expected. Initiation factor Rrn3 was only slightly enriched at the spacer promoter and highly enriched at the gene promoter and tails out in the 5' region of the rDNA gene, which was in line with its direct binding to Pol I during initiation and the dissociation of Rrn3 from Pol I after promoter escape during the transition to productive elongation phase[19]. Pol I, itself was highly enriched between spacer promoter and its corresponding termination site as well as over to whole transcribed region of the 47S rRNA gene as expected in growing cells. Ubf, the six HMG-box containing Pol I transcription factor specific to *metazoa* (see result section 2.3 and Figure 11B), was found at the spacer and the gene promoter as well as in the enhancer region, and was enriched over the whole transcribed region, showing some peaks of higher local Ubf concentration, indicating preferred Ubf-binding sites. The results for Top2a showed an enrichment over the entire rDNA gene including spacer and gene promoter as well as the enhancer repeat region and some higher concentration peaks. The ChIP-Seq pattern of Top2a clearly did not resemble the pattern of Pol I initiation factors Rrn3, Taf1b or TBP nor of Pol I itself, but was comparable to the ChIP-Seq pattern of Ubf (Figure 23A). Apparently, some of the peaks with higher local concentration of Top2a overlayed with preferred Ubf-binding sites.

As ChIP-Seq results pointed to the possibility of direct Top2a - UBF binding via the HMG-boxes of UBF, this option was further investigated. First, endogenous Top2a was pulled down from human cell lysate using a Top2a antibody and co-immunoprecipitants were analyzed. Western blot analysis showed that Pol I subunit RPA49, representing hPol I, as well as UBF were pulled down by Top2a (Figure 23B). Binding of recombinant *in vitro* purified proteins was tested to clarify whether UBF-Top2a interaction was direct. Human flag-tagged UBF (fUBF) and human Top2a were pre-incubated and later co-purified on beads with anti-Flag antibody as shown by a Western Blot against UBF and Top2a (Figure 23C). Increasing salt concentrations (lane 2-4) weakened and finally abolished the co-IP. Notably, *in situ* docking studies, comparable to the hPol I/dock II molecular modelling, using HMG-box 5 of UBF identified similar binding sites as for dock II (Figure 23D). Potential interaction surfaces were detected for the catalytic and ATPase domain of Top2a and for both Top2a states. Hence, a direct interaction of Top2a to hPol I and to UBF might be possible in human cells and *in vitro*.

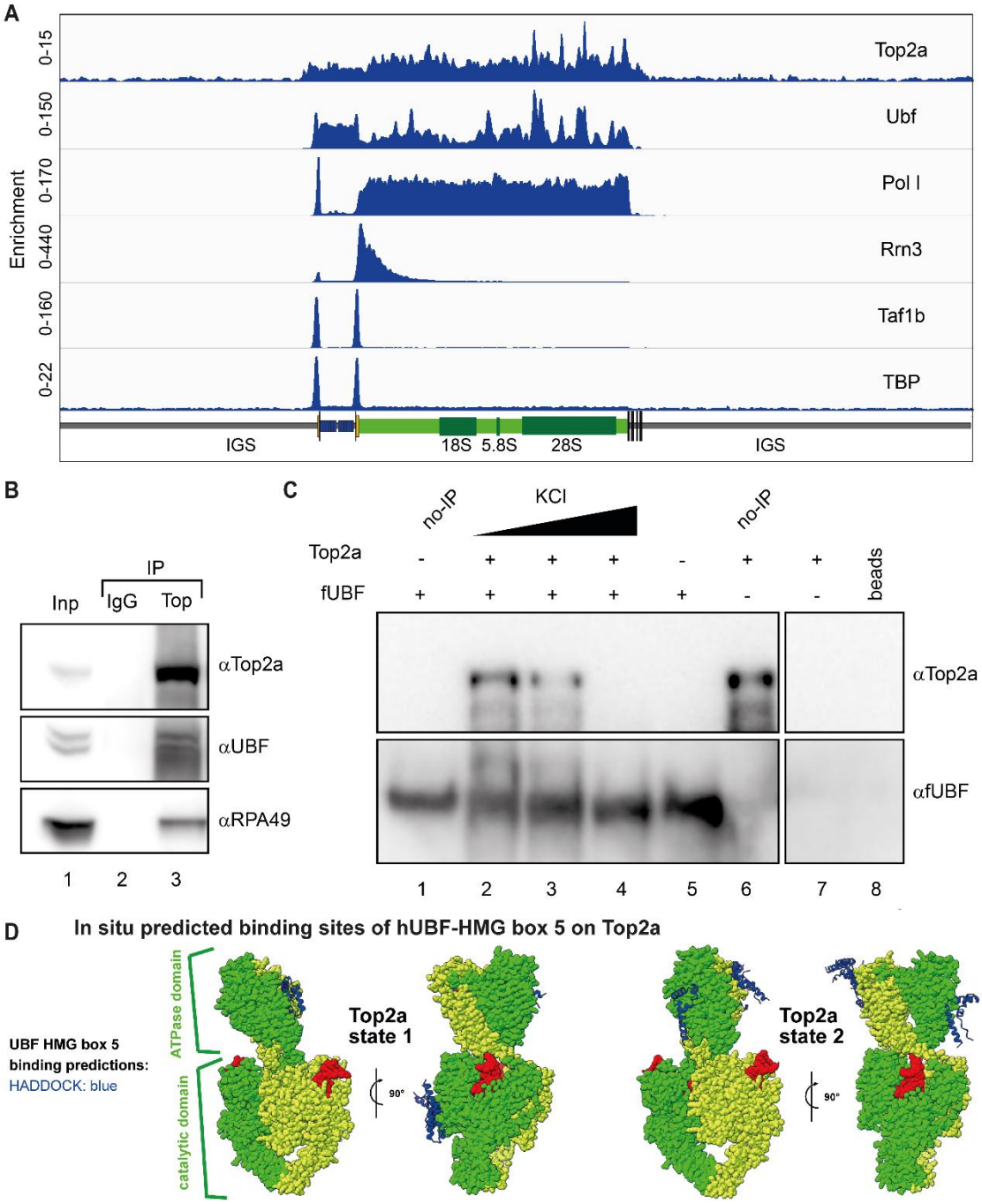


Figure 23 ChIP analysis for different factors of the Pol I system and Top2a - UBF interaction

A Top2a was detected over the entire mouse rDNA gene regions occupied by Ubf. Original raw data from reference Canela *et al.*[219] was aligned and deconvoluted as previously described[221]. Peaks over the 3' region of the gene overlapped with Ubf peaks, indicating co-localization. Top2a overlapped binding peaks for the initiation factors Rrn3, Taf1b and TBP, but specific correlations were not observed. Pol I signal marked the actively transcribed region. **B** UBF co-immunoprecipitates with Top2a: Top2a was immunoprecipitated from nuclear extract of human U2OS cells using anti-Top2a antibodies (Abcam) immobilized on magnetic beads (DynaI). Immunoprecipitated proteins were analyzed by Western blot using anti-UBF, anti-RPA49, and anti-Top2a antibodies; Lane 1: 10% input, Lane 2: IP with IgG control, Lane 3: IP with anti-Top2a antibodies. **C** Purified human Top2a co-precipitates with purified human UBF at low salt concentrations. Recombinant fUBF was incubated with purified Top2a at three different salt concentrations (lanes 2–4). Lane 1: fUBF control (no IP), Lane 5: IP without Top2a, Lane 6: Top2a control (no IP), Lane 7: IP without fUBF addition, Lane 8: FLAG-bead only control. **D** HMG-box 5 of human UBF (blue) was predicted to bind Top2a (subunits of homodimer shown in lemon and green, respectively; dsDNA in red) in states 1 and 2 (PDB 6ZY7 and 6ZY8, respectively) using ZDOCK.

3 Discussion

How hPol I evolved between yeast and human, which features are conserved or human-specific, and how hPol I adapted to the increased regulatory demands in human cells were the key questions of this thesis. A detailed structural and functional analysis was performed to investigate the human enzyme and to compare the results with the knowledge of the yeast counterpart. For this purpose, an *in vitro* system needed to be established. A human cell line with Pol I specifically and endogenously tagged was created to allow affinity purification of active human Pol I enzymes. The detailed *in vitro* structure-function analysis of hPol I now revealed insights into the evolution of Pol I, conserved features between yeast and human and human-specific details. Additionally to this work, two studies followed similar[222] or divergent[223] approaches supporting many findings and supplementing each other by focusing on different key aspects[222–224].

The HeLa-POLR1A cell line enabled a detailed hPol I localization study over the cell cycle without any antibodies. It could be shown that during interphase hPol I localizes within the nucleolus. Within this compartment, hPol I is not evenly distributed but rather accumulated within specific single spots. Throughout mitosis, these spots of high hPol I density disperse during pro- and prometaphase and hPol I is evenly distributed until the hPol I spots start to reappear in late telophase (Figure 7, see result section 2.1.3). Previously, it has been shown that Pol I distribution within the nucleolus is punctuated rather than even distributed in different mammalian cells including humans[181]. A previous study based on antibody-antigen reaction proposed that during mitosis high amounts of Pol I travel to the daughter cells bound to the DNA at the nucleolus organizer regions[181]. However, antibody-antigen reactions are not completely bias-free and could be potentially influenced by post-translational modifications of the target protein. In 2020, Maiser *et al.* found a possible explanation for the punctuated behavior of Pol I within the fibrillar centers of the nucleolus[225]. They described that single, active rDNA gene units form a ring-shaped loop of about 170 nm diameter in humans separated from each other[225]. Remarkably, they found overlapping UBF and RPA190 antibody signals at active rDNA genes[225]. A similar approach as ours, not based on antibodies, showed a clustering of hPol I during interphase as well[183]. They used a cell line with an endogenous N-terminal Dendra2-tag at subunit RPA190, in which stable hPol I clusters represent active rDNA transcription in super-resolution live cell imaging[183]. During mitosis they show that most Pol I molecules dissociate from the DNA and especially no novel DNA binding events of Pol I occur resulting in a DNA-independent distribution of hPol I between the two daughter cells[183], being in line with our observations.

In vitro structure-function analysis of Pol I was limited to yeast organisms *S. cerevisiae* and *S. pombe* until recently. Hence, a specific emphasis was put on the evolutionary conservation of Pol I. As no homologue for the yeast subunit A14 was identified using sequence-based searches on DNA and protein levels, the existence of this subunit was challenged[17]. At the same time as two other groups, we could now verify that yeast subunit A14 is missing in the human enzyme which indeed contains a single-subunit stalk (Figure 8, Figure 9, and Figure 11, see result sections 2.2 and 2.3)[222–224]. In line with our findings, that subunit A14 is present only in the class of *Saccharomycotina* within the kingdom of fungi, Misiaszek *et al.* showed that A14 is only found in some *Ascomycota*, which is the division to

which the *Saccharomycotina* belongs to[222]. Apparently, an N-terminal region of human Pol I subunit RPA43 (aa 28-44) partially takes over the position of A14 in yeast Pol I[222]. In yeast, this N-terminal portion of RPA43 interacts with a small region of subunit RPA2 (aa 1135-1168), which co-evolved with subunit A14[222]. As a result, the single-subunit stalk subcomplex has a smaller interaction surface with the polymerase core and is therefore more flexibly linked[222,223]. The 13-subunit Pol I conformation with a single-subunit stalk is the predominant configuration of Pol I throughout evolution and model organisms *S. cerevisiae* and *S. pombe* with a 14-subunit Pol I are essentially outliers in evolution.

Investigations of Pol I from different organisms have already shown that the association and dissociation of the heterodimer RPA49/34 is another feature varying among organisms[60,66], even though the yeast subcomplex remains associated over the entire rRNA gene during transcription[59,226]. Pol I purifications from yeast show that quantitative heterodimer association is stable over ion-exchange chromatography[184], but it is not stable in high salt conditions, equal to mouse Pol I[65,227]. For rat Pol I, Hannan *et al.* have shown that less than 60 % of affinity purified Pol I molecules contain subunit RPA49 concluding that different Pol I populations varying in heterodimer association may exist[60,68]. In fission yeast *S. pombe*, structural analysis of elongating and monomeric Pol I shows a flexible or even dissociated RPA49/34 heterodimer, while both subunits are more well-defined in inactive Pol I dimers[44]. Considering that RPA49/34 association/dissociation rate is different from organism to organism, these two subunits were classified as Pol I subunits (as in yeast and human) or Pol I-associated factors (as in mouse)[65]. Cross-species heterodimer building of RPA49 and RPA34 from yeast and human, respectively, is not possible, but the human heterodimer RPA49/34 was found to co-immunoprecipitate in parts with budding yeast *S. cerevisiae* Pol I[65]. Additionally, we could show that heterodimers of both organisms (*S. cerevisiae* and *H. sapiens*) functionally stimulate hPol I ΔRPA49/34 to hPol I activity in *in vitro* elongation and cleavage assays[224] (Figure 14, see result sections 2.4 and 2.5.3).

Comparing yeast and human Pol I *in vitro* activity revealed a reduced cleavage or backtracking and proof-reading activity of the human enzyme (Figure 13, see result section 2.5.2). The reason for this is unclear, but one obvious difference of the human enzyme compared to yeast is the enhanced clamp/stalk flexibility (Figure 15 and Figure 17, see result section 2.6.1), which might explain at least in parts the impaired proof-reading and cleavage activity. For *S. cerevisiae* Pol I, it was recently shown that it is more error-prone than its Pol II counterpart[186]. Yeast Pol I is also faster in transcription possessing a higher nucleotide addition rate, whereas its elongation complexes are less stable than the ones of Pol II[186]. Comparing the elongation speed of yeast and human Pol I shows that human Pol I transcribes faster (60 nt/s compared to 91 nt/s, respectively)[65,223,228]. These observations are in line with our findings of the *in vitro* transcription assays and raises the question about the error rate within rRNAs in humans. Nevertheless, it is imaginable that there might be additional, but still unknown factors which help human Pol I in proof-reading and/or cleavage of the rRNA during transcription.

Single-particle cryo-EM reconstructions of hPol I determined within this work and by two other groups in different functional states, help to understand the traits of the human enzyme and to dissect conserved from organism-specific features[222–224]. We reported the structure of monomeric hPol I[224] (Figure 17, see result section 2.6.1), while Misiaszek *et al.* determined an elongation complex (EC) in post-translocated state, an open complex (OC) (hPol I bound to an transcription bubble without RNA)

and a co-structure with initiation factor RRN3[222]. In addition, Zhao *et al.* report reconstructions of elongating polymerase in pre-translocated, post-translocated and backtracked states from overexpressed subunits[223].

Analyzing all structures revealed that the clamp and stalk region of the human enzyme is rather flexible, compared to the yeast homologues. As a consequence, technical challenges arose which were overcome with different strategies: Misiaszek *et al.* used focused refinement and classification to improve the density of the clamp/stalk region, allowing modelling of these parts using a subset of ~ 5 % of high-quality particles[222]. Additionally, the clamp of the EC was further stabilized by the presence of an unnatural dsRNA in the RNA exit tunnel[222]. Zhao *et al.* used a different strategy and performed model building with a crosslinked sample, which showed improved density for the camp/stalk region[223]. For monomeric hPol I, neither strategy resulted in improved cryo-EM densities of flexible regions[224], which lead to the conclusion that the human Pol I clamp/stalk region inherently shows a reduced rigidity and does not adopt a defined state in the absence of substrate compared to its yeast counterpart. The C-terminal region of the stalk subunit RPA43 (at least aa 205-end), containing the connector and parts of the OB domain, remained flexible in all reconstructions[222–224]. Overall, the stalk as subcomplex itself is flexibly linked to the hPol I core as it shows a reduced interaction surface and lacks contacts to the clamp domain in comparison to its yeast counterpart[222,223]. These clamp contacts in yeast are enabled by A135 (aa 1133-1168) and a N-terminal helix in the common Pol subunit Rpb6 (aa 52-68)[223], which were found to be organism-specific. Both regions are only present in the organisms possessing subunit A14[222]. The part of the human stalk subunit RPA43 (aa 28-44), which is homologous to the part performing these clamp contacts in yeast, is located differently in the human enzyme and occupies here the space of subunit A14 in yeast[222]. Hence, clamp and stalk are particularly separated in the human enzyme, leading to or co-inciting with a more closed clamp in the human enyzme[223]. In line with these findings, the hPol I EC shows increased contacts between the clamp domain and engaged DNA contributing to a tightly closed clamp and apparently resulting in elongation complexes with an increased stability[223]. During initiation and upon RRN3 binding, a 'flipping' of the stalk was observed by a ~ 5 Å movement of the distal part of RPA43, whereas the polymerase-proximal part of RPA43 remains stably anchored to the Pol I core[222]. An enhanced flexibility within the clamp/stalk region was previously described in *S. pombe* Pol I reconstructions compared to *S. cerevisiae* ones, but is even more pronounced in the human enzyme[44].

The movement of the stalk/clamp region is essential for proper Pol I activity, especially during Pol I activation by contraction[5,44,45]. The now available structures of the human enzyme in different functional states revealed that cleft contraction upon activation is a conserved feature throughout evolution (Table 4). Similar to yeast, maximal cleft contraction is observed during elongation, a slight widening takes place during backtracking[223], probably to partially loosen DNA-contacts. Monomeric hPol I showed major flexibilities excluding the observation of cryo-EM density for the clamp/stalk region. This finding indicated that this region may not adapt a defined state at all. Therefore, it can be concluded that activation by contraction is a conserved feature of the rDNA transcription system important for regulation and efficiency[148].

In line with a conserved activation strategy of Pol I, the ordering of the bridge helix and trigger loop and the detachment of the DNA mimicking expander element takes place between monomeric and EC structures of Sc and Sp Pol I[5,44]. The structures of hPol I now indicate a similar mechanism in the human system, again suggesting conservation throughout evolution[222,223]. The backtracked state of hPol I

shows a slightly extended bridge helix, in line with a partial inactivation by cleft expansion upon backtracking[223]. Furthermore, bridge helix and trigger loop are both disordered and the active-site aspartate residues coordinating the catalytic magnesium ion show poorly defined density in the open complex reconstruction of hPol I[222]. At the same time, the clamp/stalk region is well-defined suggesting an intermediately activated state which is again in line with the suggestion of a conserved activation by contraction hypothesis[44].

Table 4 Cleft and clamp-stalk distances in yeast and human

downstream cleft: *Hs* K203 of RPA1 to K1356 of RPA1; *Sc* G231 of A190 to K1331 of A190
upstream cleft: *Hs* L341 of RPA1 to I396 of RPA2; *Sc* L399 of A190 to I424 of A135
clamp - stalk: *Hs* R101 of RPA1 to P125 of RPA43; *Sc* R100 of A190 to P127 of A43

	H. sapiens				*S. cerevisiae*			
	PDB	Distance [Å]			PDB	Distance [Å]		
		downstream cleft	upstream cleft	clamp - stalk		downstream cleft	upstream cleft	clamp - stalk
Monomer	8A43				5M3M	40.9	47.6	63.4
+ RRN3	7OBA	38.5	46.2	64.1	5G5L	40.6	47.0	63.1
OC	7OBB	39.1	46.7	65.1	5M5W	39.6	45.4	64.2
EC^{post}	7VBB	31.4	35.9	69.4	5M3F	33.9	39.3	63.8
EC^{bt}	7VBC	31.4	37.4	68.3				

The C-terminal ribbon of RPA12, which shows homologies to TFIIS of the Pol II transcription system, can invade the active center through the funnel domain in order to catalyze the cleavage of backtracked RNA[52,223]. The RPA12 C-ribbon domain was shown to be flexible in pre- and post-translocated EC states, but inserted into the funnel in the backtracked state, in an open complex reconstruction and the co-structure of hPol I with RRN3[222,223]. In a backtracked state of the human enzyme which shows an inserted RPA12 C-ribbon, the funnel is slightly wider and the gating tyrosine (Y687) is opened to allow backtracking of the RNA[223]. The flipping of the gating tyrosine is induced by the insertion of RPA12 and not by backtracked RNA as can be seen by the inactive open complex state without any RNA, in which the gating tyrosine is opened as well[222].

The heterodimer RPA49/34 was shown to be a composed element which is conserved among organisms but also contains organism-specific features[56]. A divergent C-terminal tail of RPA34 is flexible in all single particle cryo-EM reconstructions of human Pol I, similar to its yeast counterpart in subunit A34.5[42,43,61,222–224]. The comparable part of subunit RPC5 in human Pol III is flexible in cryo-EM structures as well, but was shown to include two consecutive, structured tandem winged helix domains and to play a role in whole hPol III complex stability[179]. The ARM region of RPA34, connecting the dimerization domain and the flexible C-terminal tail, is divergent from yeast to human, but binds the core polymerase in the same region and possesses the same function of helping to anchor RPA34 to the core[222–224]. In yeast, contacts between the Pol I core and subunit A34.5 are mostly formed by charged residues and hydrogen bonds. Contrary, these contacts are formed by two contacting hydrophobic patches and charged residues of the RPA34 ARM domain are oriented in a distal direction in the human counterpart and may function as an additional interaction platform[222]. The N-terminal dimerization domains of RPA34 and RPA49 are visible in all reconstructions, despite showing rather weak density which may indicate a flexible association[222–224]. The linker, helix-turn-helix motif, and tWH

domains of RPA49 are flexible in the monomeric state in both, human and yeast, enzymes[224]. In the inactive OC-state and bound to RRN3, these regions were shown to be flexible in the human enzyme[222,223]. Intense focused classification and refinement of cryo-EM data allowed modelling of the region in a small subset of particles in a post-translocated EC[222]. Despite low conservation between yeast and human, RPA49 is structurally conserved with the linker spanning over the cleft and the tWH binding the clamp near the RNA exit channel[222]. Superimposing the different structures of the human enzyme revealed that the tWH helix would clash with bound RRN3, which indicates a role in assisting RRN3-release during promoter escape for the tWH domain[222].
On top of the 'activation by contraction' mechanism for Pol I, 'hibernation by dimerization' was found to play a role *in vitro* and in yeast cells as suggested by structural and biophysical studies from *S. cerevisiae* to *S. pombe*[44]. For both organisms, Pol I homodimerization can take place, but is achieved by divergent elements. This raises the question whether the mechanism of dimerization might be conserved throughout evolution. *In vitro* analyses of human Pol I did not yield indications for dimerization in metazoans. While this observation does not exclude the possibility of dimerization in a salt- and concentration-dependent manner similar to yeast, but with the equilibrium shifted to the monomeric population, current evidences suggest that Pol I dimerization may be exclusive to yeast species[44].

The most striking observation in human Pol I structures was the discovery of a previously undescribed structural domain within its largest subunit RPA1. This domain was called 'dock II' and resembles an HMG-box fold situated on the downstream side of the polymerase (Figure 19, see result section 2.6.1)[224]. *In situ* three-dimensional fold analysis revealed that the domain is closely related to structures of HMG-box 5 within the hPol I transcription factor UBF[224]. Unlike most HMG-boxes, the dock II domain is incapable of DNA-binding, but may serve as a protein-protein interaction surface (Figure 20, Figure 21, Figure 22, and Figure 23, see result sections 2.7.1 and 2.7.2). Dock II is located within the foot domain, which is known to be a protein-protein interaction hotspot in Pol II[204,205,229]. Mammalian foot domain of Pol II was also shown to be altered compared to yeast, possibly creating an enhanced interaction platform[229]. It can be speculated, that the divergent foot of human Pols compared to yeast may be required to allow interaction with an increased number of transcription factors in metazoan cells. Accordingly, phylogenetic analysis confirms a high conservation of the domain in metazoans, but an absence in other organisms (Figure 11, see result section 2.3)[224].
The search for possible interaction partners of the human dock II domain yielded Topoisomerase 2a (Top2a) as a likely candidate (Figure 21 and Figure 22, see result section 2.7.2). Topoisomerases type II, controlling DNA topology in an ATP-dependent manner via double-strand breaks, are expressed as two isoforms in vertebrates, Top2a and Top2b, whereas Top2a is expressed in proliferating cells and Top2b throughout the cell cycle[230,231]. Top2a was previously shown to be involved in hPol I transcription initation[210]. The deletion of the C-terminal domain (CTD) of Top2a leads to a six-fold reduction of Top2a co-purification with RRN3, but only a two-fold reduction with hPol I co-purification[210], suggesting that the CTD promotes the binding to RRN3, but that different Top2a regions may interact with hPol I directly. Furthermore, Top2a can be stimulated by, and interacts with an HMG-box domain of the protein of HMGB1 which does not bind to DNA as well[211]. In line with the observations that Top2a is part of the hPol I transcription system and can interact with HMG-box domains, we found that dock II domain might transiently and directly interact with Top2a (Figure 21, see result section 2.7.2). *In*

situ molecular dynamic simulations supported the possibility of direct interaction between Top2a and hPol I (Figure 22, see result section 2.7.2). Promoting a role of Top2a in rDNA transcription, an enrichment of Top2a at the rDNA locus was uncovered by ChIP-Seq analysis in mouse cells (Figure 23, see result section 2.7.2). Unlike transcription initiation factors, Top2a apparently bound the entire promoter and transcribed region. Some Top2a peaks co-localized with the six HMG-boxes containing factor UBF and a physical interactions between Top2a and Pol I, as well as Top2a and UBF were shown (Figure 23, see result section 2.7.2). Interestingly, it remains unclear what the exact functions of the HMG-boxes within UBF are. UBF is conserved throughout metazoans, indicating co-evolution with the 'dock II' domain (Figure 11, see result section 2.3). *In situ* calculations have shown that HMG-box 5 of UBF can potentially bind Top2a in a similar way as dock II domain (Figure 23, see result section 2.7.2). Moreover, Top2a can pull down UBF and hPol I from cell lysates, indicating an interaction in a similar way as dock II domain (Figure 23, see result section 2.7.2).

These conclusions strengthen the notion that Top2a plays a role within the Pol I transcription system in metazoans. Nevertheless, it remains unclear what is its specific function. Speculations suggest three possible scenarios (Figure 24): (1) Top2a may assist Pol I initiation by releasing supercoiling tension during DNA melting, (2) Top2a may behave similar to a transcription elongation factor with the goal of releasing DNA-supercoiling tension to aid subsequent Pol I passage, especially during the first round of transcription, or (3) Top2a may bind to UBF, constructing hubs of increased Top2a concentrations all over the gene with the possibility of being transiently handed over to hPol I whenever necessary[224]. The finding, that Top2a is part of hPol I initiation complexes supports scenario (1) and would allow to release the tension which accumulates upon spontaneous DNA melting during initiation[210]. In the Pol II initiation system, a similar mechanism is found: the XPB translocase of TFIIH is located in a similar position in yeast and human and functions to release tension during DNA melting[232–234]. However, Top2a ChIP-Seq data did not confirm a specific accumulation of Top2a on the Pol I promoter. Scenario (2) is supported by the fact that positive supercoiling always accumulates in the direction of transcription, which is even augmented for Pol I due to the higher loading rate and speed compared to other polymerases[65,138,235,236]. Especially, during the pioneering round of transcription of previously inactive genes, release of supercoiling tension could be important for proper transcription. To release the supercoiling, Top2a might bind to the downstream edge of hPol I via the dock II domain similar to an elongation factor. Scenario (3) is based on the association of Top2a to UBF, which may create periodic Top2a hubs, which allows the transient recruitment and handing over of Top2a to hPol I on active genes to release supercoiling tension[224]. As additional supercoiling is introduced by the high on/off rate of UBF in actively transcribed genes, the third scenario may provide an elegant mechanism to deal with both reasons for supercoiling: UBF binding and hPol I transcription. Overall, the *in vitro*, *in situ*, and in cell analysis support scenario (3) the most, but does not exclude any of the other scenarios and a combination of the three potential scenarios is possible as well.

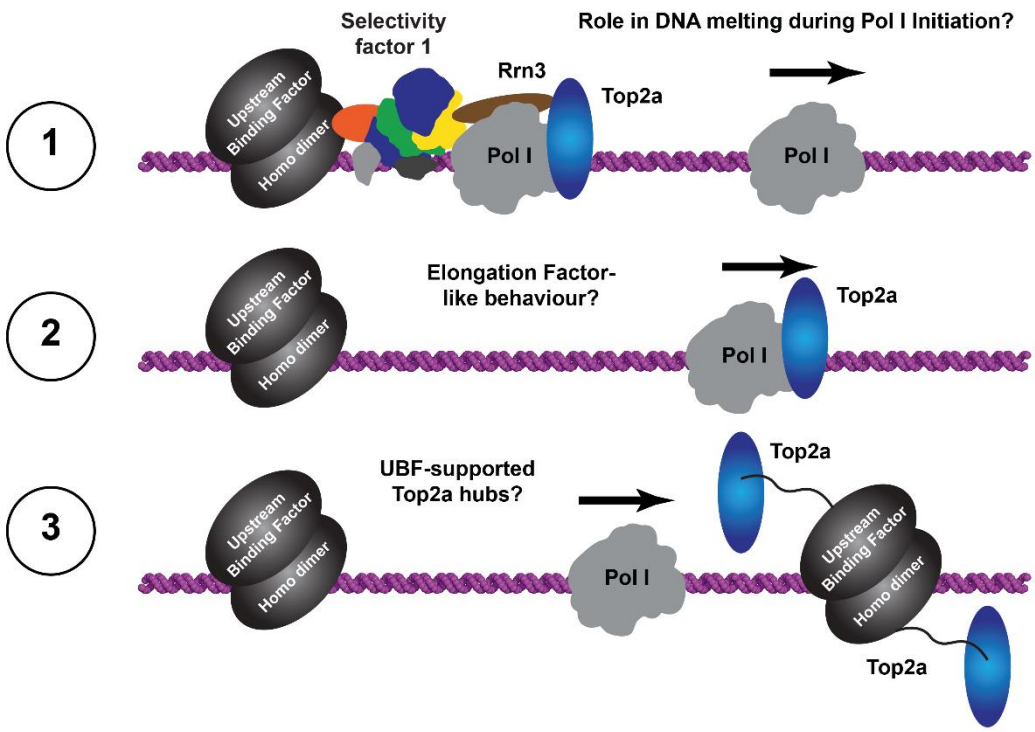


Figure 24 Three scenarios of Top2a action during Pol I transcription

Three hypotheses are likely scenarios: (1) Top2a may support initiation by resolving supercoils generated during dsDNA melting. (2) Top2a may travel with Pol I in an elongation factor-like manner to resolve positive supercoils upon their accumulation. (3) UBF and Top2a cooperate to form "torsion release hubs" at the 3' region of the rDNA gene and Top2a can be handed over transiently to hPol I.

Top2a and Top2b were shown to be in an equilibrium between nucleoplasm and nucleolus, being highly mobile[231,237,238]. ATP depletion leads to an accumulation of Top2a and Top2b within the nucleolus[231,237]. Interestingly, Pol I inhibition by the small molecule BMH-21 abolishes nucleolar localization of both Top2 isoforms even under ATP depletion[231]. Additionally, BMH-21 treatment disrupts nucleolar localization of Pol I transcription factor UBF associating to active rDNA genes in human cells[231]. BMH-21 is a selective inhibitor of Pol I transcription by intercalating in GC-rich DNA regions, particularly in G-quadruplexes[231,239,240]. Hence, BMH-21 slows down nucleotide addition and thus elongation speed, induces long-living pausing, and reduces Pol I occupancy on the rDNA[241–243].

With the cryo-EM reconstructions of human Pol I determined, structures of all three DNA-dependent RNA polymerase from yeast and mammals are now available (Figure 25). All six Pols are structurally resolved in various functional states. This allowed a detailed analysis of structural and functional conservation among species and across transcription systems, as well as the prediction of the influences disease-causing mutations may have on the enzymes. In general, the architecture of the polymerases is conserved between yeast and human. One difference applying for all three Pol structures was a divergent foot domain comparing yeast and human enzymes. Particularly, hPol I comprised the additional dock II domain, which might serve as an additional protein-protein interaction platform. The clamp/stalk flexibility seemed to be enhanced in human apo Pol I and hPol I possessed a single-subunit

stalk, which is the pre-dominant configuration across species. Moreover, hPol I, compared to its yeast counterpart, showed a reduced cleavage and proof-reading activity *in vitro*.

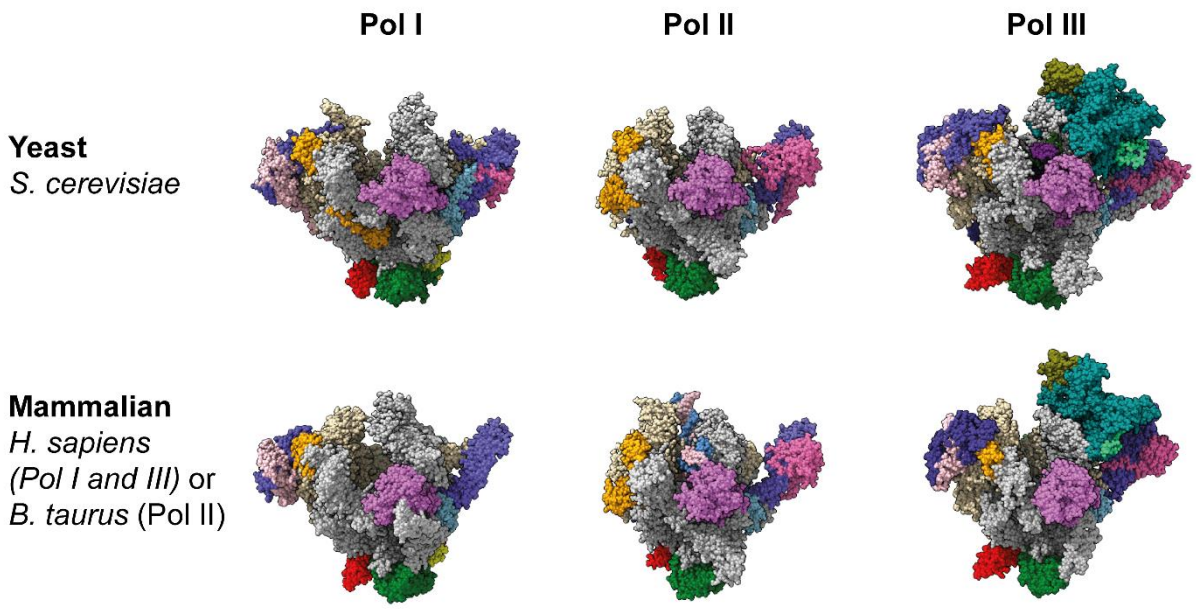


Figure 25 Structures of Pol I, Pol II and Pol III of yeast and mammals.

Front view of Pol I, II, and III from Saccharomyces cerevisiae (top row; PDB 4C2M, 1WCM, 5FJ8) and from mammals (bottom row; this study, PDB 5FLM, 7AST (associated to this study)). Subunits as in Table 1; Color code as in Fig 2: RPA1/RPB1/RPC1: grey; RPA2/RPB2/RPC2: wheat; RPAC1/RPB3: red; RPAC19/RPB11: yellow; RPABC1: magenta; RPABC2: hafnium; RPABC3: green; RPABC4: lemon; RPABC5: density; RPA43/RPB7/RPC8: slate; RPA12/RPB9/RPC10: orange; RPA49/RPC5: light blue; RPA34/RPC4: pink; A14/RPB4/RPC9: hot pink; RPC3: cyan; RPC6: olive green; RPC7: light green.

4 Outlook

The availability of the structures in different functional states in combination with *in vitro* functional analysis can now support the basic understanding of hPol I transcription and in the next step of disease mechanisms. These findings now lay the groundwork for further research and drug development for Pol I associated diseases. While the key concepts of the Pol I activity cycle and its regulatory mechanisms appear to be conserved, the detailed roles and structural basis of regulation by organism-specific factors in response to external or cell-dependent queues remain poorly understood. How human Pol I initiation, elongation, backtracking, proof-reading, termination, and chromatin interaction are achieved, what the exact functions of mammalian-specific domains are, and how mutations in the Pol I system exactly contribute to human diseases will be key questions over the following years. Some of the questions can now be addressed on the basis of recent analysis, such as the nature of and reason for the apparently decreased proof-reading efficiency and the increased nucleotide-incorporation error-rates of human compared to yeast Pol I. However, disease mechanisms and the study of the direct and indirect implications that Pol I subunits and rRNA-related transcription factors and their mutants have, will be more difficult to understand. Especially the precise role of Top2a in interplay with the human Pol I transcription system, function(s) of the metazoan dock II domain and implications of the absence of stalk-subunit A14 in humans will remain challenging to study. Answers to these questions hold the promise to give insights into understanding why the Pol I transcription system, with its enormous specialization to the synthesis of a single product type, is of such importance to all eukaryotes.

5 Material

5.1 Strains

BL21 (DE3) RIL	*E. coli*	F– ompT hsdS(rB– mB–) dcm+ Tetr gal λ(DE3) endA Hte [argU ileYleuW Camr]	Agilent
BL21 (DE3) R3 pRARE	*E. coli*	Derived from BL21(DE3) and Rosetta2 (Merck). BL21(DE3) transformed with plasmid pRARE2 (isolated from Rosetta2 calls), which carries seven rare-codon tRNA genes.	SGC
BL21 (DE3) pLysS	*E. coli*	F–, ompT, hsdSB (rB–, mB–), dcm, gal, l(DE3), pLysS, Cmr	Progema
DH5alpha	*E. coli*	F- _80lacZ_M15 _(_lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 _	Invitrogen
XL1blue	*E. coli*	recA1, endA1, gyrA96 thi-1, hsdR17, supE44, relA1, lac [F´proAB lacIqZΔM15 Tn10 (TetR)]	Stratagene
DH10EmBacY	*E. coli*		Geneva Biotech, Imre Berger[244]
pirHC	*E. coli*		Geneva Biotech, Imre Berger[244]
pirLC	*E. coli*		Geneva Biotech, Imre Berger[244]
SF21	*Spodoptera frugiperda*		Imre Berger[244]
HeLa P2	uterus; human adenocarcinoma		Philip Gunkel[180]
HeLa-POLR1C	uterus; human adenocarcinoma	CRISPR/Cas9 RPAC1-GS.3C-sfGFP	this study
HeLa-POLR1A	uterus; human adenocarcinoma	CRISPR/Cas9 RPA1-GS.3C-sfGFP	this study
HEK293T	kidney; human embryonic cells		ATCC (CRL-3216)
HEK-POLR1A	kidney; human embryonic cells	CRISPR/Cas9 RPA1-GS.3C-sfGFP	established by Andreas Schmidbauer
Flp-In T-REx-293	kidney; human embryonic cells	T-Rex/Flip-In System	Invitrogen/Thermo Fisher Scientific
T-Rex-293-RPA2	kidney; human embryonic cells	T-Rex/Flip-In System RPA2-GS.3C-sfGFP-Flag-HA	this study
T-Rex-293-RPA43	kidney; human embryonic cells	T-Rex/Flip-In System RPA43-GS.3C-sfGFP-Flag-HA	this study
T-Rex-293-RPA34	kidney; human embryonic cells	T-Rex/Flip-In System RPA34-GS.3C-sfGFP-Flag-HA	this study
T-Rex-293-TAF1A	kidney; human embryonic cells	T-Rex/Flip-In System TAF1A-GS.3C-sfGFP-Flag-HA	this study

5.2 Media

Bacteria	LB	1 % (w/v) NaCl 1 % (w/v) Bacto Tryptone 0.5 % (w/v) Yeast Extract + 2 % (w/v) Agar for plates
	TB	1.2 % (w/v) Bacto Tryptone 2.4 % (w/v) Yeast Extract 0.5 % (v/v) Glycerol 2x Salt solution NPS
	SOB	10 mM NaCl 2.5 mM KCl 2 % (w/v) Bacto Tryptone 0.5 % (w/v) Yeast Extract 10 mM $MgCl_2$ 10 mM $MgSo_4$
	Salt solution NPS 20x	0.5 M $(NH_4)_2SO_4$ 1 M KH_2PO_4 1 M Na_2hPO_4
	5052 mix	25 % (w/v) glycerol 2.5 % (w/v) glucose 10 % (w/v) α-Lactose
	Antibiotics (final concentrations for bacterial selection)	Ampicillin: 0.1 mg/ml Kanamycin: 0.05 mg/ml Chloramphenicol: 0.03 mg/ml Spectinomycin: 0.05 mg/ml Tetracyclin: 0.05 mg/ml Streptomycin: 0.01 mg/ml Gentamycin: 0.01 mg/ml X-Gal: 0.025 mg/ml
Insect cells	Sf-900 II SFM	Gibco/Thermo Fisher Scientific
Human cells	DMEM media	Gibco/Thermo Fisher Scientific
	DMEM high glucose	Gibco/Thermo Fisher Scientific
	FBS	Gibco/Thermo Fisher Scientific
	Pen/Strep	Gibco/Thermo Fisher Scientific
	PBS	Gibco/Thermo Fisher Scientific
	Trypsin	Gibco/Thermo Fisher Scientific
	Antibiotics (final concentrations)	Hygromycin B 100 µg/ml (Invitrogen/Thermo Fisher Scientific) Zeocin 100 µg/ml (Invitrogen/Thermo Fisher Scientific) Blasticidin 5 µg/ml (Invitrogen/Thermo Fisher Scientific)
	Tetracyclin	1 µg/ml (Invitrogen/Thermo Fisher Scientific)

5.3 Oligonucleotides

ID	name	sequence	method
	EC_39nt_t	AAGCTCAAGTACTTAAGCCTGGTCATTACTAGTACTGCC	cryo-EM
	EC_39nt_nt	GGCAGTACTAGTAAACTAGTATTGAAAGTACTTGAGCTT	cryo-EM
	EC_LNA	UAUCUGCAUGUAGACCAGGC	cryo-EM
JD64	EC5-templatestrand	CGAGGTCGAGCGTTGTCCTGGT	trxn assay
JD65	EC5-nontemplatestrand	CGCTCGACCTCG	trxn assay
JD66	EC5-RNA-Fam	5'-Fam-AACGGAGACCAGGAC	trxn assay
JD324	EC5-bubble-NTS	CTGATCGTCGTAGCTCACACTGTCCGCTCGACCTCG	trxn assay
JD325	EC5-bubble-TS	CGAGGTCGAGCGTTGTCCTGGTGCTACGACGATCAG	trxn assay
JD320	EC3-templatestrand	CGAGGTCGAGCGTGTCCTGGTCTAG	trxn assay
JD321	EC3-RNA-Fam	5'-Fam-GACCAGGAC	trxn assay
JD322	EC6-templatestrand	CGAGGTCGAGCGTTGTCCTGGTCTCCGT	trxn assay
JD323	EC6-RNA-Fam	5'-Fam-ACGAGACGGAGACCAGGAC	trxn assay
	EC_x-template-DNA	AAGTCAAGTACTTACGCCTGGTCATTACTAGTACTGCC	trxn assay
	EC_x-tail-nontemplate-DNA	TAGTACTTGACTT	trxn assay
	EC_x-bubble-nontemplate-DNA	GGCAGTACTAGTAAACTAGTATTGAAAGTACTTGACTT	trxn assay
	EC_x-RNA-Cy5	5'-Cy5-UAUAUGCAUAAAGACCAGGC	trxn assay
2207	Tailed template competitor	CGAGTAAGTATAGGGTAAGGTGAT	tail template
4019	3' r bio 601 2k	GAAGCCATACCAAACGACGAGC	tail template
4220	tail fwd Kompetitor rev	ATCACCTTACCCTATACTTACTCG	tail template
	EMSA-DNA-strand1	5'-Cy5- CTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGA	EMSA
	EMSA-DNA-strand2	TCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAG	EMSA
JD433	hsPromoter_-312_fwd_pOPIN-B	AAGTTCTGTTTCAGGGCCCGGTGTGTCCCGGTCGTAGGAG	cloning/ PCR for EMSA
JD434	hsPromoter_+100_rev_pOPIN-B	ATGGTCTAGAAAGCTTTAAGCGCGGCCGGCTAG	cloning/ PCR for EMSA
JD03	A190 gRNA fwd strand1	CACCGGCTCCAAGGACCCTTGGTGA	cloning
JD04	A190 gRNA fwd strand2	AAACTCACCAAGGGTCCTTGGAGCC	cloning
JD05	A190 gRNA rev strand1	CACCGGCGGGGTAGCTGCTATCTCAG	cloning
JD06	A190 gRNA rev strand2	AAACCTGAGATAGCAGCTACCCCGCC	cloning
	pX-U6-XbaI-fwd	AATGCTCTAGAGAGGGCCTATTTCCCATGATTCC	cloning
	pX-U6-KpnI-rev	TGCATGGTACCGCCATTTGTCTGCAGAATTGGCGC	cloning
JD01	A190-Homologie1_fwd-Primer +EcoRI	GACTGAATTCAGAGATGCCCCAGACCACG	cloning
JD02	A190-Homologie2_rev-Primer +HindIII	GATCAAGCTTGTCCACACCAAACTGACAGCAG	cloning
JD07	A190-Homologie-linker-sfGFP first half 40bp overhang	* gBlock	cloning
JD08	A190-Homologie-linker-sfGFP second half 40bp overhang	* gBlock	cloning
JD38	A34-fwd.FL(1-)_pOPINE	AAGAGATATACCATGGAGGAGCCCCAGGC	cloning
JD39	A34-fwd.10-_pOPINE	AAGAGATATACCATGGCTCGGTTCTCTTGTCCCC	cloning
JD40	A34-fwd.29-_pOPINE	AAGAGATATACCATGCGTTTCTCCTTGGAGGCG	cloning

JD41	A34-fwd.131-_pOPINE	AAGAGATATACCATGCAGCCCATCCCAGCAAGTC	cloning
JD42	A34-rev.-122_pOPINE	GTGATGGTGATGTTTTTACTGGGGACCCTCAAGGATC	cloning
JD43	A34-rev.-134_pOPINE	GTGATGGTGATGTTTTTATGGGATGGGCTGCAGAG	cloning
JD342	HsA34-rev.-151_pOPIN-E	GTG ATG GTG ATG TTT TTA GGCACAGAACCGAGGC	cloning
JD44	A34-rev.FL-510_pOPINE	GTGATGGTGATGTTTTTACACAGGCTGCTGCTGC	cloning
JD45	A49-fwd.FL(1-)_pOPINB	AAGTTCTGTTTCAGGGCCCGATGGCGGCGGAGGTG	cloning
JD46	A49-rev.-115_pOPINB	ATGGTCTAGAAAGCTTTAATCTGAAAATAGTGGCTGCATGTTG	cloning
JD47	A49-fwd.180-_pOPINB	AAGTTCTGTTTCAGGGCCCGGGTGTGACTGCTCTGGTCAG	cloning
JD48	A49-fwd.206-_pOPINB	AAGTTCTGTTTCAGGGCCCGGATGCAGCCAAGCCTGAAG	cloning
JD49	A49-rev.-404_pOPINB	ATGGTCTAGAAAGCTTTAAGGCAGCGGGAGGG	cloning
JD50	A49-rev.FL(-420)_pOPINB	ATGGTCTAGAAAGCTTTACAAGGTAATCTTCCTCCGCTTTGC	cloning
JD310	A190.1060-1155_CodonOPt_ pOPIN-B	* gBlock	cloning
JD311	A190.1081-fwd_pOPIN-B	AAGTTCTGTTTCAGGGCCCGATGAATACTCTGCTTCGTCGCG	cloning
JD312	A190.1146-rev_pOPIN-B	ATGGTCTAGAAAGCTTTAACTCGGATCTGGACAGGC	cloning
JD332	HsA12-fwd_pOPIN-B	AAGTTCTGTTTCAGGGCCCGATGTCTGTCATGGACCTCGC	cloning
JD333	HsA12-rev_pOPIN-B	ATGGTCTAGAAAGCTTTAAGAGTCTTCCTTCTCCTGGAACTTG	cloning
JD336	ScA12-fwd_pOPIN-B	AAGTTCTGTTTCAGGGCCCGATGTCTGTTGTAGGATCGTTAATTT TTTGCTTGG	cloning
JD337	ScA12-rev_pOPIN-B	ATGGTCTAGAAAGCTTTAATTGTTGGTACGGAACTTGTAACCAC AG	cloning
JD10	Rrn3-pOPINB/M-fwd	AAGTTCTGTTTCAGGGCCCGATGGCAGCCCCGC	cloning
JD11	Rrn3-pOPINB/M-rev	ATGGTCTAGAAAGCTTTACAATGGAGACGGCTGC	cloning
JD15	hsRrn3-Sf21optimized	* gBlock	cloning
JD78	His-hsRrn3optSF21_Infusion_ BamHI_fwd	CCGTCCCACCATCGGGCGCGGGACATGGGCAGCAGCCATCATC	cloning
JD79	His-MBP-hsRrn3optSF21_Infusion _BamHI_fwd	CCGTCCCACCATCGGGCGCGGGACATGGCACACCATCACCACC	cloning
JD80	hsRrn3optSF21_Infusion_SalI_rev	TTACAATGGAGACGGCTGCATGGGACGCGGCCGCGACAAGTGA GCTCGTC	cloning
JD347	HsUBF-fwd.1-_pOPIN-B	AAGTTCTGTTTCAGGGCCCGATGAACGGAGAAGCCGACTG	cloning
JD348	HsUBF-rev.-764(FL)_pOPIN-B	ATGGTCTAGAAAGCTTTAGTTGGAGTCAGAGTCTGAGGAG	cloning
JD28	5xGSfront-fwd	GGCTCCGGTTCAGGTTCTG	cloning
JD308	sfGFP-rev_-Stop_Inf.pcDNA5	GGCGGCGGATCCGAGCTCTCCCTTGTACAGCTCGTCCATGCC	cloning
JD302	A135-fwd_Inf.pcDNA5	CTCTAGCGTTTAAACTTGCGATGAGAGCAAAAGAGGCTGC	cloning
JD58	A135-rev+5xGS-linker	CCGCTGCCACTGCCAGAACCTGAACCGGAGCCAACAACATCCAG TTTCACTTTGATGTTC	cloning
JD304	A43-fwd_Inf.pcDNA5	CTCTAGCGTTTAAACTTGCGATGGCTGCAGGTTGCTCAG	cloning
JD21	A43-rev+5xGS-linker	ACCGCTGCCACTGCCAGAACCTGAACCGGAGCCAAGAAAATTA CTTTTCCCTTTTCTTTTTGGTGAG	cloning
JD305	A34-fwd_Inf.pcDNA5	CTCTAGCGTTTAAACTTGCGATGGAGGAGCCCCAGGC	cloning
JD24	A34-rev+5xGS-linker	ACCGCTGCCACTGCCAGAACCTGAACCGGAGCCCACAGGCTGCT GCTGCTG	cloning
JD271	TAF1A_fwd+KpnI	AGCTGGTACCATGAGTGATTTCAGTGAAGAATTAAAAGGGC	cloning
JD27	TAF1A-rev+5xGS-linker	ACCGCTGCCACTGCCAGAACCTGAACCGGAGCCGAGTCTTGGAT TTACAATACTGTATTTTTTCACAGATC	cloning

* Sequences of gBlocks:

JD07_A190 Homologie-linker-sfGFP first half 40bp overhang:

AGATGCCCCAGACCACGGGGAAGGGTTAAGTAGCGTGAGCAGGTGGCGGAGCTGAAGGGGAGCCCCAGAGTCTAACTCCCAAGCCATGTGCTGATGTG
GAGGCAGCTTGGCTTGTGTCTGCTGACTGCTCCCGTGGTCACTGCTTAATCCCTCCTAATTTTACCTCCCGTTTGACAGGCATCGCGGTCGACCCTCGCCATC
TCTCCCTGGTTGCTGATTATATGTGCTTCGAGGGTGTTTACAAGCCACTGAATCGCTTTGGGATCCGGTCAAACTCTTCCCCGCTACAGCAGATGACATTTG
AAACCAGCTTCCAGTTTCTGAAGCAAGCCACCATGCTGGGTTAGTGCGTGGGCAGGAGCCCTGGCCTCCCCTCTACTGACGACTCTGGGCACACGGGCTCA
CCCATGGGCATCTTTCCATGGTTCCTTGCACACCAGTGGGATCTTGGGATCTGAGGGGAGCGGGGGCGGCTGGAGGGGGCTGAACAGCAGACACGACAG
CTGCCTCTACGCTTTGAAGGAACAGGAAGAGGGCCTTCCCCTCTCAAGTGCTAGAGGGGGGCTGCCTCGTGATCCCATCAGCTTCAGAGACTCAGGACCCA
CCTGAGTCCCGAGACTTCAGTGTTCAGCCCCATAATCACCTCATGACATTCCCAACTTCTACCTCAACACCCTGACACATGTTTTTGTGTCTCTGTCAGGATCC
CACGATGAGCTGAGGTCTCCTTCTGCCTGCCTTGTGGTCGGGAAGGTCGTCAGGGGCGGGACAGGCCTGTTCGAGCTCAAGCAGcctCTGAGAGGCTCCG
GTTCAGGTTCTGGCAGTGGCAGCGGTCTCGAAGTCTTGTTCCAAGGGCCTGGCTCTGGGTCAGGAAGCGTGAGCAAGGGCGAGGAGCTGTTCACCGGGG
TGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGCGCGGCGAGGGCGAGGGCGATGCCACCAACGGCAAGCTGACCC
TGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACA
TGAAGCGCCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCAGCTTCAAGGACGACGGCAACCTACAAGACCCGCGCCGA
GGTGAAG

JD08_A190 Homologie-linker-sfGFP second half 40bp overhang:

CAAGGACGACGGCACCTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG
ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTTCAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAA
GATCCGCCACAACGTGGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCACCCAGTCCGTGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCACGGCA
TGGACGAGCTGTACAAGTAGCAGCTAcccCGGCACCATCTGCCCAGCTCCAAGGACCCTTGGTGAgcgCGTGGCCCAGCCTGCCTTCTGCATGAGAGGACCA
GGAGACTGGAATCCAGGGCAGTTCCAAGTGACAGTACAGAGCACAGCAGCGACCTTGGGCCTGAAAGCAGTGGGCCTCTGAGCTGGGCCAGCTTCACCT
GGAAAGTGACAGAGTTGCTCATCCTTGCCCCTCCCTGTCTCTGGATTTTTATCAAGGTTTACCAAGTCTTCTGAGTCCCCCTGAGATGGCTGGGGCCTCACCT
GTGCTGCAGGAGGCCTCTGTGGCATAACCCCTAAGGAGAAGTCCTGATTCACGATTCACTGAGAAGACCAAGGGGAAGCCATGCTTTGCTGCTGGGGACC
CCAGGCACCTCCAGAGTAGGGAAGCGGGGTTCTTTTGCTGTGAGTGGCCAGGGACAACAGACAAGATTCCTGGGGGCTCCCGATGAGCAGGAACGTGGA
GCCTGCTGCCCAAGGCCTGCTCCTTCCGGCTGCTCCAGCCCCTGGGGGCAGAGTCCACAAAGAGTCCCCATCAAGACTTCTTCCCTGAGTCAAGTACAGCG
TAGCATAGTCCTCCACCCACCCAACCTCTCTGCCTGGCCAGGGTCCTGGCCCTGCCACTGTGTGGCGAGGTGTCCTTCTAGACCACATCAGCCCCAAGGCTG
GGAGCAGTCGCTCCAGGGCCGCAGCAGTTCACTCCCACACATAGAACCCAGGTCACTGCTGGGGCGATTGAACAGGTTGCCTGGCTTTTCTCTGCTGTCAG
TTTGGTGTGG

JD310_A190.1060-1155_CodonOPt_pOPIN-B:

AAGTTCTGTTTCAGGGCCCGGCAGACCCGAAGAAGGCCCTGCATCACTTCCGCGCCATCAAGAAGTGGCAGTCAAAACATCCAAATACTCTGCTTCGTCGC
GGTGCTTTTTTGTCCTACTCACAAAAGATTCAAGAAGCTGTGAAAGCCCTGAAGTTAGAATCTGAAAACCGTAACGGTCGCTCTCCCGGCACACAAGAGAT
GCTTCGCATGTGGTATGAATTGGACGAAGAATCTCGCCGCAAATACCAGAAGAAGGCAGCCGCCTGTCCAGATCCGAGTCTTAGTGTATGGCGCCCCGAT
ATCTACTAAAGCTTTCTAGACCAT

JD15_hsRrn3-Sf21optimized:

ATGGCAGCCCCGCTCTTGCACACGAGATTGCCCGGAGATGCAGCTGCGAGCTCGTCTGCAGTCAAGAAACTGGGTGCTTCACGCACCGGTATTTCCAATAT
GAGGGCGCTTGAAAATGATTTTTTTAATTCGCCACCACGCAAGACTGTAAGATTCGGTGGAACTGTAACGGAGGTACTGCTGAAGTATAAAAAGGGTGAG
ACCAATGACTTCGAACTCCTTAAAAATCAACTTCTCGATCCGGACATAAAGGACGACCAAATAATTAATTGGTTGCTTGAATTCAGATCGAGCATAATGTAC
CTCACTAAGGATTTTGAGCAGCTGATATCTATAATACTGCGTTTGCCTTGGCTGAATCGCAGCCAAACGGTGGTTGAAGAATACCTGGCATTCCTGGGTAAC
TTGGTAAGTGCACAGACGGTATTTTTGAGACCCTGTTTGTCTATGATTGCCAGCCATTTCGTGCCACCACGTGTAATCATTAAAGAAGGTGACGTAGATGTG
TCAGATTCTGATGATGAGGACGATAATCTCCCGGCGAACTTCGATACATGCCATCGTGCGCTTCAAATCATTGCACGTTATGTGCCATCAACTCCCTGGTTTT
TGATGCCTATTCTCGTAGAAAAATTTCCCTTCGTACGTAAATCGGAGAGGACCTTGGAATGCTATGTTCATAACCTCCTCCGTATATCTGTGTACTTCCCCAC
ACTTAGACATGAAATTCTGGAGCTTATCATCGAAAAGTTGCTCAAGTTGGACGTAAACGCATCGCGCCAGGGCATCGAGGATGCAGAAGAGACGGCGACA
CAGACCTGTGGCGGAACCGATAGTACGGAGGGCTTGTTCAATATGGATGAAGATGAGGAAACGGAACACGAGACCAAAGCGGGACCGGAGAGGCTCGA
TCAGATGGTCCACCCCGTCGCAGAGAGACTGGATATTCTTATGTCTTTGGTACTTTCTTATATGAAAGACGTTTGTTACGTGGACGGAAAGGTCGATAACG
GAAAGACAAAAGACCTCTACAGGGATCTGATTAATATATTCGATAAACTGCTTTTGCCCACACATGCATCGTGTCATGTGCAGTTCTTCATGTTTTACCTGTG
TTCATTTAAGTTGGGCTTTGCAGAGGCATTTCTGGAACATCTCTGGAAAAAGTTGCAGGACCCTTCTAATCCGGCCATTATCCGCCAGGCCGCGGGTAATTA
TATTGGCTCCTTCCTCGCAAGGGCCAAGTTCATTCCTCTCATAACAGTCAAATCCTGTCTCGATCTGCTGGTGAACTGGTTGCACATCTACCTTAATAACCAG
GATTCCGGAACCAAGGCGTTTTGTGACGTTGCCCTCCATGGTCCGTTTTATAGTGCCTGCCAAGCCGTCTTTTATACTTTTGTTTTTAGGCACAAACAGCTGC
TGTCAGGCAACCTTAAGGAAGGATTGCAGTATCTTCAGTCTTTGAATTTCGAGCGCATAGTCATGAGTCAACTGAACCCGCTGAAAATTTGTCTCCCCAGTG
TGGTGAACTTTTTTGCCGCCATAACAAACAAGTACCAGCTCGTGTTCTGTTATACTATCATAGAAAGGAATAACAGACAGATGCTTCCTGTAATTAGAAGCA
CCGCCGGCGGCGACTCCGTCCAGATATGTACGAACCCGCTCGACACATTTTTTCCTTTTGACCCATGCGTCCTGAAGCGTAGTAAGAAATTCATAGACCCGA
TCTATCAAGTGTGGGAGGACATGAGTGCTGAAGAGCTTCAAGAATTCAAGAAACCCATGAAGAAAGATATCGTCGAGGATGAGGACGACGACTTCTTGAA
AGGTGAGGTACCCCAGAATGATACAGTAATCGGCATAACTCCATCCTCATTCGACACTCATTTTAGAAGTCCCAGTAGCTCTGTTGGTTCACCGCCGGTTCT
TTACATGCAGCCGTCTCCATTG

5.4 Plasmids

ID	name	origin	method
#2318	pUC19 tail g- ext w/o BS	Department of Biochemistry III, University of Regensburg	tail template
424	pUC9 HrDNA1	Department of Biochemistry III, University of Regensburg	
374	pOPIN-B_hsPolIPromoter_-312-+100	this study	PCR/EMSA
	pUC19_hsPolIPromoter_-312-+100	this study	PCR/EMSA
194	pX462 (pSpCas9n(BB)-2A-Puro)	Addgene/Feng Zhang[245]	CRISPR/Cas9
294	pX462_POLR1A-gRNAs-fwd+rev	this study	CRISPR/Cas9
198	pET28b-POLR1A-Homologie-GS-linker.3C-sfGFP	Philip Gunkel	CRISPR/Cas9
214	pOPIN-B	Christoph Engel (Cramer Lab MPI Göttingen)	*E. coli* expression
377	pOPIN-B_A190.1060-1155	this study	*E. coli* expression
380	pOPIN-B_A190.1081-1146	this study	*E. coli* expression
360	pOPIN-B_hsA34.152-510	this study	*E. coli* expression
361	pOPIN-B_hsA34.186-510	this study	*E. coli* expression
236	pOPIN-B_hsA49-FL	this study	*E. coli* expression
239	pOPIN-B_hsA49.1-115	this study	*E. coli* expression
237	pOPIN-B_hsA49.180-404	this study	*E. coli* expression
238	pOPIN-B_hsA49.209-404	this study	*E. coli* expression
370	pOPIN-B_HsRPA12-FL	this study	*E. coli* expression
372	pOPIN-B_ScA12.2-FL	this study	*E. coli* expression
217	pOPIN-B_HsRRN3-FL	this study	*E. coli* expression
340	pOPIN-B_UBF1-FL	this study	*E. coli* expression
341	pOPIN-B_UBF2-FL	this study	*E. coli* expression
369	pOPIN-B_TOP2A-ATPase.1-431	this study	*E. coli* expression
366	pOPIN-B_TOP2A-cat.431-1217	this study	*E. coli* expression
368	pOPIN-B_TOP2A-ΔCTD.1-1217	this study	*E. coli* expression
367	pOPIN-B_TOP2A-FL	this study	*E. coli* expression
216	pOPIN-M	Christoph Engel (Cramer Lab MPI Göttingen)	*E. coli* expression
378	pOPIN-M_A190.1060-1155	this study	*E. coli* expression
381	pOPIN-M_A190.1081-1146	this study	*E. coli* expression
371	pOPIN-M_HsRPA12-FL	this study	*E. coli* expression
373	pOPIN-M_ScA12.2-FL	this study	*E. coli* expression
218	pOPIN-M_HsRRN3-FL	this study	*E. coli* expression
365	pOPIN-M_TOP2A-cat.431-1217	this study	*E. coli* expression
215	pOPIN-E	Christoph Engel (Cramer Lab MPI Göttingen)	*E. coli* expression
240	pOPIN-E_hsA34-FL	this study	*E. coli* expression
458	pOPIN-E_hsA34.1-343	this study	*E. coli* expression
241	pOPIN-E_hsA34.1-134	this study	*E. coli* expression
242	pOPIN-E_hsA34.10-134	this study	*E. coli* expression
243	pOPIN-E_hsA34.29-122	this study	*E. coli* expression
244	pOPIN-E_hsA34.131-510	this study	*E. coli* expression
199	pACEBac1	Geneva Biotech (Imre Berger)	SF21 expression
219	pACEBac1_His-HsRRN3-FL	this study	SF21 expression
220	pACEBac1_His-MBP-HsRRN3-FL	this study	SF21 expression

250	pcDNA5 FRT/TO	Invitrogen/Thermo Fisher Scientific	T-REx system
252	pcDNA5 FRT/TO - FLAG-HA	Invitrogen/Thermo Fisher Scientific	T-REx system
348	pcDNA5 FRT/TO_A135-GS.3C-sfGFP-FLAG-HA	this study	T-REx system
347	pcDNA5 FRT/TO_A43-GS.3C-sfGFP-FLAG-HA	this study	T-REx system
346	pcDNA5 FRT/TO_A134-GS.3C-sfGFP-FLAG-HA	this study	T-REx system
349	pcDNA5 FRT/TO_TAF1A-GS.3C-sfGFP-FLAG-HA	this study	T-REx system
251	pOG44	Invitrogen/Thermo Fisher Scientific	T-REx system
353	VP5	Department of Biochemistry I, University of Regensburg; Roeder lab, New York[246]	expression in human cell culture
24	pET28b_ScRrn3	Claudia Blattner[89]/Christoph Engel	*E. coli* expression
23	pET28b_ScA49/34.5	Sebastian Geiger[56]/Christoph Engel	*E. coli* expression
362	pVE-VV_TOP2A	Valerie Lamour	
364	pET28b_TOP2A-cat	Valerie Lamour	
363	pAX3_TOP2A-CTD	Valerie Lamour	

5.5 Antibodies

antibody	species	dilution	source	method
α-mouse 800	goat, monoclonal	1 : 10 000	Li-COR (Lincoln, USA)	Western Blot
α-His	mouse, monoclonal	1 : 1 000	Quiagen: 34660	Western Blot
α-GFP	mouse, monoclonal	1 : 1 000	Chromotek	Western Blot
α-RPA1	mouse, monoclonal	1 : 100	Santa Cruz: sc-48385	Western Blot
α-RPAC1	mouse, monoclonal	1 : 100	Santa Cruz: sc-374443	Western Blot
GFP-booster AlexaFlour 488	alpaca, recombinant nanobody	1 : 500	Chromotek	Immunoflourescence
anti-Top2a antibody	rabbit, polyclonal		Abcam: ab12318	IP; Western Blot
anti-UBF	rabbit, polyclonal		Santa Cruz: sc-9131	Western Blot
anti-RPA49	mouse, monoclonal	1 : 1 000	BD Transduction: 611413	Western Blot

5.6 Buffers

general solutions	TBS-T	10 mM Tris pH 7.2 - 7.4 150 mM NaCl 0.05 % (v/v) Tween-20
	TBE	90 mM Tris 90 mM Boric acid 1 mM EDTA
	PBS	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 1.8 mM KH_2PO_4
	TE	10 mM Tris pH 8.0 1 mM EDTA
	MOPS running buffer	50 mM MOPS 50 mM Tris base 0.1 % (w/v) SDS

		1 mM EDTA → pH ≈ 7.7
	WB transfer buffer	25 mM Tris pH 8.5 192 mM glycerin 20 % (v/v) methanol
	WB blocking buffer	5 % (w/v) milk powder in TBS-T
	WB mild stripping buffer	200 mM Glycine 0.1 % (w/v) SDS 1 % (v/v) Tween 20
	Proteinase K buffer	20 mM Tris pH 7.5 300 mM NaCl 25 mM EDTA 2 % (w/v) SDS
	IF wash buffer	0.05 % (v/v) Triton X-100 1 % (w/v) BSA in PBS
	Protein-Protein binding buffer	20 mM Hepes pH 8.0 150 mM NaCl 50 mM KCl 1 mM $MgCl_2$ 2 % (v/v) glycerol 2 mM β-mercaptoethanol
	100x Protease Inhibitor mix (PI)	100 mM PMSF 200 mM Benzamidin in 100% EtOH
	5x SDS loading dye	130 mM Tris pH 6.8 15 % (w/v) glycerol 2.1 % (w/v) SDS 8,4 % (v/v) β-mercaptoethanol 0.15 % (w/v) bromophenol blue
	6x NP OrangeG loading dye	10 mM Tris pH 7.6 60 mM EDTA 60 % glycerol 0.03 % Orange G
	2x RNA loading dye	8 M Urea 2x TBE 0.03 % (w/v) bromophenol blue 0.03 % (w/v) xylene cyanol
	2x RNA loading dye (tail template)	8 M Urea 2x TBE 10 mM EDTA 10 mM EGTA 0.03 % (w/v) bromophenol blue 0.03 % (w/v) xylene cyanol
gels	SDS-PAGE	separation gel: 375 mM Tris pH 8.8 0.1 % (w/v) SDS 12 % (w/v) Arcylamide (37.5:1) 0.1 % (w/v) APS (stock 10 %) 0.1 % (v/v) TEMED

		stacking gel: 125 mM Tris pH 6.8 0.1 % (w/v) SDS 3,9 % (w/v) Arcylamide (37.5:1) 0.1 % (w/v) APS (stock 10 %) 0.1 % (v/v) TEMED
	Agarose gel	1x TBE 0.8 - 2 % (w/v) Agarose 1x SYBR Safe DNA gel stain (Invitrogen/Thermo Fisher Scientific)
	native PAGE (EMSA)	10 % native PAGE: 0.4x TBE 10 % (w/v) Arcylamide (37.5:1) 0.08 % (w/v) APS (stock 10 %) 0.08 % (v/v) TEMED 6 % native PAGE: 0.4x TBE 6 % (w/v) Arcylamide (37.5:1) 0.08 % (w/v) APS (stock 10 %) 0.08 % (v/v) TEMED
	Urea PAGE gel	Elongation/Cleavage assay: 7 M Urea 20 % (w/v) Acrylamide (19:1) 1x TBE 0.05 % (w/v) APS (stock 10 %) 0.1 % (v/v) TEMED Tail template assay: 7 M Urea 6 % (w/v) Arcylamide (37.5:1) 1x TBE 0.08 % (w/v) APS (stock 10 %) 0.1 % (v/v) TEMED
Silver staining	Fixation Solution	50 % (v/v) methanol 12 % (v/v) acetic acid 0.02 % (v/v) formaldehyde
	VL Solution	0.8 mM $Na_2S_2O_3$
	Staining Solution	12 mM $AgNO_3$ 0.03 % (v/v) formaldehyde
	Developing Solution	566 mM Na_2CO_3 0.04 mM $Na_2S_2O_3$ 0.02 % (v/v) formaldehyde
Competent *E. coli* cells	TfbI	30 mM KAc 50 mM $MnCl_2$ 100 mM KCl 15 % (v/v) glycerol adjust pH to 5.8 with 0.2 M acetic acid
	TfbII	10 mM MOPS 75 mM $CaCl_2$ 10 mM KCl 15 % (v/v) glycerol

		adjust pH to 7 with 10 M NaOH
	Inoue buffer	10 mM PIPES pH 6.7 55 mM $MnCl_2$ 15 mM $CaCl_2$ 250 mM KCl
hPol purification buffers	hPol Lysis Buffer	20 mM Hepes pH 7.8 420 mM NaCl 1 mM $MgCl_2$ 10 µM $ZnCl_2$ 0.5 % (v/v) NP-40 4 mM β-mercaptoethanol 1x PI
	hPol Wash Buffer	20 mM Hepes pH 7.8 420 mM NaCl 1 mM $MgCl_2$ 10 µM $ZnCl_2$ 2 % (v/v) glycerol 4 mM β-mercaptoethanol
	hPol Mono A buffer	20 mM Hepes pH 7.8 1 mM $MgCl_2$ 10 µM $ZnCl_2$ 2 % (v/v) glycerol 5 mM DTT
	hPol Mono B buffer	20 mM Hepes pH 7.8 1 M $(NH_4)_2SO_4$ 1 mM $MgCl_2$ 10 µM $ZnCl_2$ 2 % (v/v) glycerol 5 mM DTT
human RPA49/34 purification	RPA49/34 Lysis buffer	50 mM MES pH 6.3 300 mM NaCl 10 mM β-mercaptoethanol 1x PI
	RPA49/34 Wash I buffer	50 mM MES pH 6.3 1 M NaCl 10 mM β-mercaptoethanol 1x PI
	RPA49/34 ATP-Wash buffer	50 mM MES pH 6.3 1 M NaCl 10 mM β-mercaptoethanol 1x PI 2 mg/ml denatured proteins 0.5 mM ATP
	RPA49/34 Wash II buffer	50 mM MES pH 6.3 300 mM NaCl 10 mM imidazole 10 mM β-mercaptoethanol 1x PI
	RPA49/34 Elution buffer	50 mM MES pH 6.3 300 mM NaCl 200 mM imidazole 10 mM β-mercaptoethanol

		1x PI
	RPA49/34 Mono A buffer	50 mM Tris pH 7.5 10 mM β-mercaptoethanol
	RPA49/34 Mono B buffer	50 mM Tris pH 7.5 2 M NaCl 10 mM β-mercaptoethanol
	RPA49/34 SEC buffer	50 mM Tris pH 7.5 150 mM NaCl 5 mM DTT
Yeast A49/34.5 purification	A49/34.5 Lysis buffer	50 mM Tris pH 7.5 300 mM NaCl 10 mM β-mercaptoethanol 1x PI
	A49/34.5 Wash I buffer	50 mM Tris pH 7.5 1 M NaCl 10 mM β-mercaptoethanol 1x PI
	A49/34.5 Wash II buffer	50 mM Tris pH 7.5 300 mM NaCl 30 mM imidazole 10 mM β-mercaptoethanol 1x PI
	A49/34.5 Elution buffer	50 mM Tris pH 7.5 300 mM NaCl 100 mM imidazole 10 mM β-mercaptoethanol 1x PI
	A49/34.5 Dilution buffer	50 mM Tris pH 7.5 10 mM β-mercaptoethanol
	A49/34.5 Mono A buffer	50 mM Tris pH 7.5 100 mM NaCl 5 mM DTT
	A49/34.5 Mono B buffer	50 mM Tris pH 7.5 1 M NaCl 5 mM DTT
Dock II purification	Dock II Lysis buffer	50 mM MES pH 6.3 300 mM NaCl 10 mM β-mercaptoethanol 1x PI
	Dock II Wash I buffer	50 mM MES pH 6.3 1 M NaCl 10 mM β-mercaptoethanol 1x PI
	Dock II Wash II buffer	50 mM MES pH 6.3 300 mM NaCl 10 mM imidazole 10 mM β-mercaptoethanol 1x PI
	Dock II Elution buffer	50 mM MES pH 6.3 300 mM NaCl 200 mM imidazole 10 mM β-mercaptoethanol

		1x PI
	Dock II SEC buffer	50 mM Tris pH 7.5 150 mM NaCl 5 mM DTT
Human RPA12 and *S. cerevisiae* A12.2 purification	RPA12 Lysis buffer	50 mM Hepes/KOH pH 7.8 500 mM KCl 10 µM $ZnCl_2$ 10 mM imidazole 10 % (v/v) glycerol 5 mM β-mercaptoethanol 1x PI
	RPA12 Wash buffer	50 mM Hepes/KOH pH 7.8 500 mM KCl 10 µM $ZnCl_2$ 20 mM imidazole 10 % (v/v) glycerol 5 mM β-mercaptoethanol 1x PI
	RPA12 ATP-Wash buffer	50 mM Hepes/KOH pH 7.8 500 mM KCl 10 µM $ZnCl_2$ 20 mM imidazole 10 % (v/v) glycerol 5 mM β-mercaptoethanol 1x PI 2 mg/ml denatured proteins 0.5 mM ATP
	RPA12 Elution buffer	50 mM Hepes/KOH pH 7.8 500 mM KCl 10 µM $ZnCl_2$ 150 mM imidazole 10 % (v/v) glycerol 5 mM β-mercaptoethanol
	RPA12 Mono A buffer	20 mM Hepes/KOH pH 7.8 2 mM $MgCl_2$ 5 µM $ZnCl_2$ 10 % (v/v) glycerol 3 mM DTT
	RPA12 Mono B buffer	20 mM Hepes/KOH pH 7.8 2 M KAc 2 mM $MgCl_2$ 5 µM $ZnCl_2$ 10 % (v/v) glycerol 1 mM DTT
	RPA12 SEC buffer	50 mM Hepes/KOH pH 7.8 300 mM KCl 10 µM $ZnCl_2$ 10 % (v/v) glycerol 5 mM β-mercaptoethanol
Rrn3 purification from SF21 insect cells	Rrn3-SF Lysis buffer	20 mM Hepes pH 7.8 200 mM NaCl 1 mM DTT

	Rrn3-SF Elution buffer	20 mM Hepes pH 7.8 200 mM NaCl 10 mM maltose 1 mM DTT
	Rrn3-SF SEC buffer	20 mM Hepes pH 7.8 200 mM KCl 1 mM DTT
Rrn3 purification from bacteria	Rrn3 Lysis buffer	50 mM Tris-HCl pH 7.5 200 mM NaCl 10 mM imidazole 10 % (v/v) glycerol 2 mM β-mercaptoethanol 1x PI
	Rrn3 Wash I buffer	50 mM Tris-HCl pH 7.5 1 M NaCl 10 mM imidazole 10 % (v/v) glycerol 2 mM β-mercaptoethanol
	Rrn3 Wash II buffer	50 mM Tris-HCl pH 7.5 200 mM NaCl 20 mM imidazole 10 % (v/v) glycerol 2 mM β-mercaptoethanol
	Rrn3 Wash III buffer	50 mM Tris-HCl pH 7.5 1 M NaCl 20 mM imidazole 10 % (v/v) glycerol 2 mM β-mercaptoethanol 5 mM ATP
	Rrn3 Wash IV buffer	50 mM Tris-HCl pH 7.5 200 mM NaCl 10 mM imidazole 10 % (v/v) glycerol 2 mM β-mercaptoethanol
	Rrn3 Mono A buffer	20 mM Tris-HCl pH 7.5 2 mM DTT
	Rrn3 Mono B buffer	20 mM Tris-HCl pH 7.5 2 M NaCl 2 mM DTT
	Rrn3 SEC buffer	25 mM Tris-HCl pH 7.5 150 mM NaCl 2 mM DTT
SL1 purification	SL1 Lysis Buffer	20 mM Hepes pH 7.8 420 mM NaCl 1 mM $MgCl_2$ 10 µM $ZnCl_2$ 0.5 % (v/v) NP-40 5 mM ATP 4 mM β-mercaptoethanol 1x PI
	SL1 Wash Buffer	20 mM Hepes pH 7.8 420 mM NaCl

		1 mM $MgCl_2$ 10 µM $ZnCl_2$ 2 % (v/v) glycerol 4 mM β-mercaptoethanol
UBF1/2 purification	UBF Lysis buffer	20 mM Tris pH 7.5 20 mM imidazole 400 mM NaCl 1 mM DTT 1x PI
	UBF Wash I buffer	20 mM Tris pH 7.5 50 mM imidazole 400 mM NaCl 1 mM DTT
	UBF Wash II buffer	20 mM Tris pH 7.5 20 mM imidazole 1 M NaCl 1 mM DTT
	UBF ATP-Wash buffer	20 mM Tris pH 7.5 20 mM imidazole 1 M NaCl 1 mM DTT 2 mg/ml denatured proteins 0.5 mM ATP
	UBF Elution buffer	50 mM Tris pH 7.5 350 mM imidazole 250 mM NaCl 1 mM DTT
	UBF Mono A buffer	20 mM Hepes pH 7.8 2 mM $MgCl_2$ 5 µM $ZnCl_2$ 1 mM DTT
	UBF Mono B buffer	20 mM Hepes pH 7.8 2 M KAc 2 mM $MgCl_2$ 5 µM $ZnCl_2$ 1 mM DTT
	UBF SEC buffer	20 mM Tris pH 7.5 250 mM NaCl 1 mM DTT
Top2a-ATPase purification	ATPase Lysis buffer	50 mM Tris-HCl pH 8.0 500 mM NaCl 5 mM $MgCl_2$ 1x PI
	ATPase Wash buffer	50 mM Tris-HCl pH 8.0 500 mM NaCl 5 mM $MgCl_2$ 30 mM imidazole
	ATPase His-Elution buffer	50 mM Tris-HCl pH 8.0 150 mM NaCl 5 mM $MgCl_2$ 400 mM imidazole 2 mM β-mercaptoethanol

	ATPase Mono A buffer	20 mM Tris-HCl pH 8.0 5 mM $MgCl_2$ 2 mM β-mercaptoethanol
	ATPase Mono B buffer	20 mM Tris-HCl pH 8.0 1 M NaCl 5 mM $MgCl_2$ 2 mM β-mercaptoethanol
	ATPase MBP-Elution buffer	20 mM Tris-HCl pH 8.0 450 mM NaCl 5 mM $MgCl_2$ 10 mM maltose 2 mM β-mercaptoethanol
Top2a catalytic domain purification	Top2a-cat Lysis buffer	50 mM Tris-HCl pH 7.5 500 mM NaCl 10 % (v/v) glycerol 10 mM imidazole 5 mM β-mercaptoethanol 1x PI
	Top2a-cat Elution buffer	50 mM Tris-HCl pH 7.5 500 mM NaCl 10 % (v/v) glycerol 250 mM imidazole 5 mM β-mercaptoethanol
	Top2a-cat Mono A buffer	30 mM Tris-HCl pH 7.5 15 mM NaCl 1 mM EDTA 2 mM β-mercaptoethanol
	Top2a-cat Mono B buffer	30 mM Tris-HCl pH 7.5 1 M NaCl 1 mM EDTA 2 mM β-mercaptoethanol
	Top2a-cat SEC buffer	30 mM Tris-HCl pH 7.5 70 mM NaCl 1 mM EDTA 2 mM β-mercaptoethanol
EMSA	EMSA-NaCl buffer	10 mM Tris pH 7.5 50 mM NaCl 1 mM $MgCl_2$ 4 % glycerol 0.5 mM EDTA 0.5 mM DTT
	EMSA-NaCl-2 buffer	20 mM Hepes pH 7.8 150 mM NaCl 2 % glycerol 0. 2% Triton-100 0.2 % Tween-20 5 mM DTT
	EMSA-KCl buffer	10 mM Tris pH 7.5 50 mM KCl 1 mM DTT
	EMSA-$(NH_4)_2SO_4$ buffer	10 mM Hepes pH 7.8 10 mM $(NH_4)_2SO_4$

		2 % glycerol
		1 mM DTT
Transcription	Elongation/Cleavage buffer	20 mM Hepes pH 7.8
		40 mM $(NH_4)_2SO_4$
		28 mM NaCl
		8 mM $MgSO_4$
		10 µM $ZnCl_2$
		10 % (v/v) glycerol
		10 mM DTT
	TT transcription buffer	20 mM Hepes pH 7.8
		10 mM $MgCl_2$
		5 mM EGTA
		0.05 mM EDTA
		0.2 mM ATP
		0.2 mM UTP
		0.01 mM CTP
		2.5 mM DTT
	10x TT Start buffer	2 mM CTP
		2 mM GTP
		20 ng/µl Heparin
Dock II-Top2a pulldown	Dock II-Top2a binding buffer	20 mM Hepes pH 7.8
		50 mM NaCl
		1 mM $MgCl_2$
		2 % (v/v) glycerol
		0.2 % (v/v) Triton X-100
		0.2 % (v/v) Tween-20
		10 mM imidazole
		2 mM β-mercaptoethanol
	Dock II-Top2a elution buffer	20 mM Hepes pH 7.8
		50 mM NaCl
		1 mM $MgCl_2$
		2 % (v/v) glycerol
		0.2 % (v/v) Triton X-100
		0.2 % (v/v) Tween-20
		350 mM imidazole
		2 mM β-mercaptoethanol
	Dock II-Top2a high salt buffer	20 mM Hepes pH 7.8
		400 mM NaCl
		400 mM KCl
		1 mM $MgCl_2$
		2 % (v/v) glycerol
		2 mM β-mercaptoethanol
Top2a co-IP	co-IP dilution buffer	25 mM Tris-HCl pH 7.9
		12.5 mM $MgCl_2$
		10 % (v/v) glycerol
		0.03% NP40
	co-IP wash buffer	25 mM Tris-HCl, pH 7.9
		150 mM KCl
		12.5 mM $MgCl_2$
		10 % (v/v) glycerol
		0.03% NP40
UBF-Top2a pulldown	UBF-Top2a pulldown buffer	25 mM Tris-HCl pH 7.9

	12.5 mM $MgCl_2$ 10 % (v/v) glycerol 0.03% NP40

5.7 Kits

peqGOLD Plasmid Miniprep Kit II	Peqlab
Plasmid Plus Midi Kit	Quiagen
QIAquick PCR purification Kit	Qiagen
QIAquick Gel Extraction Kit	Quiagen
NucleoSpin RNA clean-up kit	Macherey-Nagel
SuperScriptII Reverse Transkriptase	Invitrogen

5.8 Enzymes

Phusion High-Fidelity DNA Polymerase	self-made or New England Biolabs (NEB)
Q5 High-Fidelity DNA Polymerase	New England Biolabs (NEB)
One*Taq* Hot Start DNA polymerase	New England Biolabs (NEB)
Restriction enzymes	New England Biolabs (NEB)
Antarctic Phosphatase	New England Biolabs (NEB)
T4 DNA ligase	New England Biolabs (NEB)
Infusion mix	Takara Bio
SuperScript II Reverse Transcriptase kit	Invitrogen/Thermo Fisher Scientific
Proteinase K	Sigma Aldrich
HRV-3C protease	self-made
TEV protease	self-made
DNase I	Promega
Benzonase	Sigma-Aldrich

5.9 Equipment

Nanodrop One	Thermo Fisher Scientific
PCR Cycler primus 25 advanced	peqLab
Thermoblock	Eppendorf
NanoDrop One	Thermo Fisher Scientific
Äkta Pure 25	GE Healthcare
Äkta Purifier	GE Healthcare
ÄKTA micro Ettan LC	Pharmacia
ÄKTA columns	HisTrap 1/5 ml (GE Healthcare) HiTrap Q HP 1 ml (GE Healthcare) HiTrap SP HP 1 ml (GE Healthcare) HiTrap Heparin 1/5 ml (GE Healthcare) MBPTrap 1/5 ml (GE Healthcare) MonoS 5/50 GL (GE Healthcare)

	MonoQ 1.6/5 PC (Pharmacia Biotech) MonoQ 5/50 GL (GE Healthcare) Superdex 75 Increase 10/300 GL (GE Healthcare) Superdex200 Increase 3.2/300 (GE Healthcare) Superdex200 Increase 10/300 (GE Healthcare) Superose6 Increase 3.2/300 GL (GE Healthcare) Superose6 Increase 10/300 (GE Healthcare)
PD10 column	GE Healthcare
JEOL 2100-F Transmission Electron Microscope operated at 200 keV and equipped with TVIPS-F416 (4k x 4k) CMOS-detector	JEOL, TVIPS
CryoArm200 cryo-electron microscope operated at 200 keV and equipped with a K2 direct electron detector	JOEL, Gatan
Clip Ring	Gatan
Clip Ring Tool	Gatan
Gatan 626 Cryo Transfer Holder	Gatan
UltiMate 3000 RSLCnano System equipped with a C18 Acclaim Pepmap100 preconcentration column (100 µm i.D. x 20 mm) and an Acclaim Pepmap100 C18 nano column (75 µm i.d. x 250 mm)	Thermo Fisher Scientifc
maXis plus UHR-QTOF System	Bruker Daltonics
Plan-Apochromat 63x/1,4 Oil DIC Objective at a Zeiss LSM980/Airyscan 2 confocal microscope	Zeiss
TwoMP Mass photometer	Refeyn
Gel Documentation Imaging System	Peqlab
Agarose gel Casting system	Bio-Rad
Agarose gel chamber	Bio-Rad
Trans-Blot Turbo Transfer System	Bio Rad
BD FACSAria™ IIu cell sorter	BD Bioscience
Screen Eraser-K	Bio-Rad
Model 583 Gel Dryer	Bio-Rad
CAWO X-Ray Cassette	CAWO
Imaging Screen-K (20x25 cm)	Bio-Rad
Odyssey Infrared Imager Model 9120	Li-COR
PMI Personal Molecular Imager FX	Bio-Rad
Typhoon FLA-9500 imager	GE Healthcare
Innova 44 Incubator Shaker Series	New Brunswick Scientific
Insect cell incubator	Binder
counting chamber Neubauer new/improved	Carl Roth
large-scale cell culture CO2 incubator	PHC
Human cell culture hood	Thermo Fisher Scientific
Cimarec Biosystem 40B controller and stirrer	Thermo Fisher Scientific
Glass Spinner Flasks (various volumes)	Corning
Branson Sonifier 450	Emerson
PowerPac HV	Bio-Rad
PowerPac Basic Power Supply	Bio-Rad
XCell SureLock Mini-Cell Electrophoresis	Invitrogen/Thermo Fisher Scientific
Milli-Q PLUS	Merck Millipore
Magnetic stirrer, MR 3000	Heidolph Instruments
Avanti J-20XP	Beckman Coulter
Centrifuge 5417R	Eppendorf

Megafuge 40R	Heraeus
Optima L-80 XP Ultracentrifuge	Beckman Coulter
Cressington 208 carbon coater	Cressington
Glow Discharge Device, Plasma Cleaner/Sterilizer PDC-3xG	Harrick
Pelco 'EasiGlow' plasma cleaner	TedPella
Typhoon FLA-9500	GE Healthcare
Vitrobot Mark IV	Thermo Fisher Scientific
XCell SureLock Mini-Cell Electrophoresis	Thermo Fisher Scientific
Flotation chamber	Mechanical workshop, University of Regensburg
- 80 °C Freezer (F570)	Eppendorf

5.10 Software

Compass 1.7 acquisition and processing software	Bruker Daltonics
Protein Scape 3.1.3	Bruker Daltonics
Mascot 2.5.1	Matrix Science
Data Analysis 4.2	Bruker Daltonics
ApE - A plasmid Editor	M. W. Davis, E. M. Jorgensen[247]
Microsoft Office (Powerpoint/Excel/Word)	Microsoft
Zeiss AxioVision	Zeiss
ZEN 3.0	Zeiss
Benchling [Biology Software] (2020)	Retrieved from https://benchling.com
AcquireMP	Refeyn
DiscoverMP	Refeyn
Fiji 2.0.0	FIJI
Imaris 9.6	Oxford Instruments
SnapGene Viewer 6.0.2	Dotmatics (snapgene.com)
SerialEM	D. N. Mastronarde[248]
RELION 3.1	S. Scheres/MRC Laboratory of Molecular Biology[190,249]
RELION 4.0	S. Scheres/MRC Laboratory of Molecular Biology[190,249]
WARP	D. Tegunov, P. Cramer[189]
Chimera 1.14rc	UCSF[250]
ChimeraX 1.3	UCSF[207]
Aline	Reference [251]
Phenix suite	Reference [193,252]
WinCoot 0.8.9.2 EL	Reference [253]
HHPRED	MPI Bioinformatics Toolkit
MODELLER	MPI Bioinformatics Toolkit
ClustalΩ	MPI Bioinformatics Toolkit
PSI-Blast	MPI Bioinformatics Toolkit
Quick2D	MPI Bioinformatics Toolkit
AlphaFold	European Bioinformatics Institute
Adobe Illustrator 2021	Adobe

5.11 Consumables

Ultrapure water (< 18 Ω) was received from a Milli-Q PLUS filtration device (Merck Millipore). Unless stated otherwise, all chemicals and solvents used in this work were purchased at the highest available purity from the following companies: Sigma Aldrich, Merck, Roth, Roche, Qiagen, Serva or J.T.Baker.

NuPAGE™ 10 % Bis-Tris gel	Invitrogen/Thermo Fisher Scientific
NuPAGE™ 12 % Bis-Tris gel	Invitrogen/Thermo Fisher Scientific
NuPAGE™ 4-12 % Bis-Tris gel	Invitrogen/Thermo Fisher Scientific
Native PAGE 3-12% gradient gel (for Blue Native)	Invitrogen/Thermo Fisher Scientific
Mini gel Casting Cassettes	Invitrogen/Thermo Fisher Scientific
DNA ladder: 1kb Plus; 1 kb; 100 bp	New England Biolabs (NEB)
6x DNA loading dye	New England Biolabs (NEB)
LDS sample buffer	Thermo Fisher Scientific
SYBR Safe DNA Gel Stain	Invitrogen/Thermo Fisher Scientific
Color Prestained Protein Standard, Broad Range	NEB (#77125, #77195)
Instant Blue	C.B.S. Scientific
SimplyBlue SafeStain	Invitrogen/Thermo Fisher Scientific
Immobilon-PSQ Transfer Membrane; PVDF	Merck Millipore Ltd.
Chromatography Paper 3mm Chr Whatman	GE Healthcare
Ni-NTA beads	Qiagen
MagneHis Ni-Particles	Progema
anti-FLAG M2 Magnetic Beads	Sigma-Aldrich
MBP beads	GE Healthcare
GFP-Trap Dynabeads	Chromotek
IgG Sepharose 6 Fast Flow	GE Healthcare
Amicon centrifugal concentrators (3/10/30/100 kDa cut-off)	Millipore
PD-10 columns	GE Healthcare/Cytiva
α-^{32}P-CTP	Hartmann Analytic
NativePAGE sample buffer	Invitrogen/Thermo Fisher Scientific
light blue cathode and anode buffer	Invitrogen/Thermo Fisher Scientific
FuGENE HD Transfection Reagent	Promega
Thymidine	Sigma-Aldrich
Prolong Gold Antifade Mountant with DAPI	Thermo Fisher Scientific
Negative stain grids (G2400C)	Plano
cryoEM grids	Quantifoil
polyethylenimine HCl Max Linear 40K	Polyscience
Rotilabo® syringe filter (0.22 µm & 0.45 µm)	Carl Roth
Mica	Plano

6 Methods

6.1 DNA manipulation

6.1.1 Plasmid purification

To isolate plasmid DNA from *E. coli*, Mini- or Midiprep kits were used according to the manufacturer's instructions. Elution was performed with MilliQ water and the concentration of plasmid DNA was determined with the Nanodrop.

6.1.2 Polymerase Chain Reaction (PCR)

Polymerase chain reactions (PCRs) were performed to amplify specific DNA pieces. A reaction mix containing 1 ng/µl DNA template, 0.5 µM of each Primer, 1x HF buffer, 0.4 mM dNTPs each and home-made Phusion polymerase was used. For GC-rich DNA pieces, Phusion polymerase (NEB) was used with 1x GC buffer and additional 6 % (v/v) DMSO. The reaction itself was performed in the PCR Thermocycler as summarized in Table 5. The resulting DNA was either purified using a PCR purification kit and checked on an agarose gel or purified via agarose gel and the corresponding Agarose gel extraction kit.

Table 5 PCR cycler protocol

Initial Denaturing	98 °C	5 min	
Denaturing	98 °C	15 s	
Annealing	58 °C	20 s	25-35 cycles
Elongation	72 °C	1 kb/min	
Final Elongation	72 °C	5 - 10 min	
Storage	8 °C	∞	

6.1.3 Restriction enzyme digestion

3-5 µg plasmid DNA or PCR product was digested with one or more restriction enzymes in their favorable buffer as indicated by the manufacturer for 1-3 h at 37 °C. In case of a vector backbone in a ligation-dependent cloning, dephosphorylation was additionally performed before the desired DNA was purified by agarose gel electrophoresis.

6.1.4 Dephosphorylation

Vector backbone DNA for a ligation-dependent cloning was dephosphorylated with Antarctic Phosphatase with its corresponding buffer for 1 h at 37 °C followed by purification over an agarose gel.

6.1.5 Ligation

Dephosphorylated vector backbone DNA and insert DNA was ligated with the help of T4 DNA Ligase in 1x T4 Ligation buffer for 1 h at room temperature and afterwards transformed in *E. coli*.

6.1.6 Infusion

Linearized vector backbone and insert DNA with suitable overhangs were combined in a reaction mix containing 50 ng vector backbone, about 3 - 4 times molar access of insert DNA and 0.5x Infusion mix. After 15 min at 50 °C the DNA was ready to transform in *E. coli*.

6.1.7 Reverse Transcription

In order to produce cDNA from extracted RNA for cloning, the SuperScript II Reverse Transcriptase kit was used with random primers according to the manufacturer's instructions.

6.1.8 Purification of PCR products

In order to separate DNA from the other components of a PCR reaction, a PCR purification kit was used as in the manufacturer's instructions. Elution was performed with MilliQ water, the concentration measured with the Nanodrop and the DNA checked on an agarose gel.

6.1.9 Purification of nucleic acid from agarose gels

Correct DNA bands were cut out from the agarose gel and purified with an Agarose gel extraction kit according to manufacturer's instructions. Elution was performed with MilliQ water and the concentration determined with the Nanodrop.

6.1.10 Agarose gel electrophoresis

DNA samples (analytical: 100 - 300 ng DNA; preparative: everything, up to 5 µg) was mixed with DNA loading dye to 1x concentration. The samples were loaded on 0.8 - 2 % (dependent on DNA product length) agarose gels containing 1x SybrSafe together with a DNA standard. The gel was run for 30 - 45 min at 150 V and DNA visualized under UV light.

6.1.11 Ethanol precipitation

In order to precipitate nucleic acids at least 300 mM NaCl (final concentration) was added to the nucleic acids and afterwards 100 % ethanol to reach 70 - 95 % (v/v) ethanol concentration (final). The reaction was inverted vigorously and incubated at -20 °C for at least 1 h. Afterwards centrifugation at 20 000 g at 4 °C for 30 min was performed. The supernatant was removed and the nucleic acid pellet washed with 70 % ethanol for DNA or 80 % ethanol for RNA samples by centrifugation at 20 000 g at 4 °C for 10 min. Before resolving the nucleic acid with water, the pellets were dried at 55 °C. In case, low amounts of RNA were precipitated 0.2 - 1 µg/µl glycogen (final concentration) was added in the first step to ensure complete precipitation.

6.1.12 Isopropanol precipitation

Genomic DNA was precipitated by isopropanol. Here, at least 300 mM NaCl in the sample was ensured before adding 43 % (v/v) isopropanol (end concentration). The sample was vigorously inverted, incubated at -20 °C for at least 30 min and centrifuged at 20 000 g for 30 min at 4 °C. The DNA pellet was washed with 70 % ethanol and centrifuged for another 10 min. Then the pellet was dried at 55 °C, before resuspended in water and stored at 4 °C.

6.1.13 Sequencing

Plasmid DNA or PCR products were sent for sequencing (Microsynth Seqlab GmbH) either with pre-mixed primer or without primer. The samples were handed according to the company's instructions. For samples without pre-mixed primers, a standard primer from the company's stock was selected. Sequencing results were analyzed by sequence alignments against the desired sequence and positively verified clones were selected.

6.2 Protein analysis

6.2.1 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-gels (12 %) were self-made, containing separation and stacking gel. Samples were mixed with SDS loading dye to 1x final concentration and loaded on the gel together with pre-stained protein ladder. Gel run was performed for 30 min at 200 V with MOPS running buffer. Alternatively, ready-to-use 12 % SDS gels or 4 - 12 % gradient gels (Invitrogen) were used and run for 60 min or 50 min, respectively, at 170 V.

6.2.2 Coomassie staining

Commonly, SDS gels were stained with Instant Blue overnight and destained with water for at least 2 h. Alternatively, for higher sensitivity SimplyBlue SafeStain was used with the 'maximum sensitivity protocol' as in the manufacturer's instructions.

6.2.3 Silver staining

In case of very low protein amounts, SDS gels were silver stained. First, the gel was fixed with fixation solution for at least 1 h and then washed for 20 min with 50 % ethanol. After incubating the gel for 1 min in VL solution, it was washed three times with water for 20 s each. Staining was performed for 20 min in staining solution before washing the gel two times with water for 20 s each. Developing solution was added to the gel and staining was performed as long as appropriate before stopping the reaction with solid citric acid.

6.2.4 Western blot analysis

In order to detect specific epitopes by antibodies, separated proteins from SDS-gels were blotted on a PVDF membrane using the semi-dry blotting method. On the anode a stack of 3 Whatman paper pre-soaked in WB transfer buffer, the membrane - first activated in methanol and then washed with WB transfer buffer, the SDS-gel and 3 pre-soaked Whatman paper were placed. The Trans-Blot Turbo Transfer System was used for blotting with the 'TurboBlot' settings and repeated one to two more times.

Complete blotting was verified by Ponceau S staining. For this purpose, the membrane was first washed with water, before incubating the membrane with Ponceau S solution for 10 min. Unspecific staining was washed away with water.

For antibody-detection the membrane was first incubated with WB blocking buffer for 1 h at room temperature or overnight at 4 °C. Primary antibody was diluted in WB blocking buffer as listed and incubated for 2 h at room temperature or overnight at 4 °C. After washing the membrane with TBS-T three times for 5 min, an appropriate secondary antibody was diluted in WB blocking buffer as indicated and incubated with the membrane for 1 h at room temperature. Before detecting the bound antibodies via its fluorescent label with the Odyssey Infrared Imager (Li-COR), the membrane was washed three times with TBS-T and once with water.

6.2.5 Protein quantification

The concentration of purified protein samples was determined with the Nanodrop based on the specific molecular weight and extinction factor of the protein, calculated by the ExPASy protparam web server[254]. hPol I concentration was determined by quantifying Coomassie-stained band intensity of hPol I RPA1 and RPA2 compared to titrated *S. cerevisiae* Pol I A190 and A135.

6.3 Work with *E. coli*

6.3.1 Cultivation

E. coli were cultivated in liquid LB media or on LB-agar plates at 37 °C unless otherwise stated. Appropriate antibiotic was added to the media in case of selection for plasmid transformation, as well as Chloramphenicol for expression strains. Growth of bacteria in liquid media was tracked by optical density measured at 600 nm (OD_{600}).

6.3.2 Preparation of competent *E. coli* cells

A 50 ml pre-culture of the appropriate *E. coli* strain was grown over night at 37 °C in SOB media and used to inoculate the 200 ml SOB media main culture to an OD_{600} of 0.2. The main culture was harvested once it reached an OD_{600} of 0.5 by centrifugation at 2 500 g for 10 min at 4 °C. The cell pellet was resuspended in 15 ml TfbI and incubated for 20 min. After another centrifugation step, the pellet was resuspended in 4 ml TfbII, incubated for 10 min, aliquoted in 50 µl, flash-frozen in liquid nitrogen and stored at -80 °C.

In order to transform two or more plasmids in one heat shock transformation, ultra-competent *E. coli* cells were prepared. For this purpose, 25 ml of SOB media was inoculated with one colony of appropriate *E. coli* strain and grown for 6 - 8 h at 37 °C. With 0.5 ml of this culture a second pre-culture of 125 ml SOB media was inoculated and grown over night at 18 °C to an OD_{600} of about 1.0. On the next day, 125 ml fresh SOB media was inoculated to an OD_{600} of 0.1 and grown to an OD_{600} of 0.55. Before harvesting the cells by centrifugation at 2 500 g for 10 min at 4 °C, the culture was cooled down for 10 min in 0 °C ice bath. The cell pellet was once resuspended and washed with Inoue buffer, before resuspending the cells in 10 ml Inoue buffer supplemented with 7.5 % (v/v) DMSO. After incubating the cells on ice for 10 min, they were aliquoted to 50 µl, flash-frozen in liquid nitrogen and stored at - 80 °C.

6.3.3 Heat shock transformation

Competent *E. coli* cells were thawed on ice and 10 - 100 ng plasmid DNA or the complete ligation or Infusion reaction was added to the cells before incubating them for 30 min on ice. Bacteria took up the DNA during a 1 min heat shock at 42 °C and were placed on ice for 5 min. To establish antibiotic resistance encoded on the introduced plasmid, cells were incubated with 500 µl LB media at 37 °C for 1 h, before plating them on LB-agar plates with appropriate selection antibiotic. After overnight incubation at 37 °C single colonies were picked.

6.4 Work with insect cells

For recombinant protein expression in SF21 insect cells, the MultiBac system, developed by Imre Berger[244,255], was used (Geneva Biotech, v5.1).

6.4.1 Cultivation

SF21 cells were cultivated adherently or in suspension in SF-900 II medium at 27 °C and 100 rpm. In suspension, cell density was kept constantly in the range of $0.5*10^6$ - $2*10^6$ cells/ml. For splitting, cells were diluted with fresh media to $0.5*10^6$ cells/ml.

6.4.2 Bacmid preparation

ORFs of the desired gene including the tag were cloned into pACE vector backbones. Transforming these plasmids in DH10EmBacY *E. coli* cells lead to the integration of the expression cassette into the viral bacmid genome using a transposase (Tn7). After heat shock, a long regeneration time without antibiotics was executed for 6 - 16 h, before plating the cells on LB plates with antibiotics and suitable for blue/white screening. White colonies were used for bacmid isolation and verified to be white by streaking on a new plate.

For bacmid preparation the solution of the Miniprep kit (PeqLab) were used. The cell pellet was resuspended in 300 µl solution I and mixed with 300 µl solution II. After 3 min of incubation, solution III was added, and the sample inverted. Two steps of centrifugation (13 000 g, 10 min) enabled complete removement of the white precipitate and the supernatant was precipitated with 700 µl of isopropanol.

After washing with 70 % of ethanol, the pellet was dried and gently resuspended in 30 µl of water under the sterile hood.

6.4.3 Virus generation and amplification

For virus generation, the bacmid was transfected into SF21 cells with FuGene transfectant. Hence, 100 µl SF-900 II medium was mixed well with 10 µl FuGene transfectant and incubated for 30 min. Afterwards, 20 µl of the purified bacmid solution was mixed with 200 µl SF-900 II medium and 100 µl of the medium-FuGene-mix, before incubating for 15 - 30 min. For transfection, $1*10^6$ cells were seeded into a 6 well plate and allowed to adhere for 15 - 30 min before removing the medium and adding 3 ml of fresh medium as well as 160 µl of transfection mix (bacmid/medium/FuGene mix). Cells were incubated for 48 h at 27 °C and 100 rpm. Before collecting the supernatant of the cells (V_0) containing the virus, YFP expression was checked. To amplify the virus, 50 ml of SF21 cells with a density of $0.5*10^6$ cells/ml were infected with 25 - 200 µl of V_0 virus stock. Cells should now double at least once, but amplification should also not exceed 5 (to maximum 7) days. Hence, cells were counted every 24 h and once cells reached over $1*10^6$ cells/ml, cells were diluted to $0.5*10^6$ cells/ml and grown for another 24 h. The supernatant (V_1) was stored at 4 °C for large-scale infections and the cell pellet can be used for a test purification if applicable.

6.4.4 Large-scale infection of insect cells and recombinant protein expression

Large-scale infections of SF21 were performed to allow for recombinant protein expression. Hence, 1 L of SF21 cells with $0.5*10^6$ cells/ml (within a 5 L Erlenmeyer flask) were infected with 5 - 20 ml V_1 virus stock. Protein expression was executed for 48 - 60 h and cells harvested by centrifugation. Cell pellets were flash-frozen in liquid nitrogen and stored at -80 °C.

6.5 Work with human cell culture

6.5.1 Cultivation

Adherent cell lines (HeLa, HEK, T-REx-293 and their derivates) were cultivated in DMEM media supplemented with 10 % FBS and 1 % Penicillin/Streptomycin at 37 °C and 5 % CO_2 atmosphere. Media of T-REx-293 cells was additionally supplemented with 100 µg/ml Zeocin and 5 µg/ml Blasticidin and derivates of T-REx-293 with a gene already specifically inserted for overexpression were grown with 100 µg/ml Hygromycin B and 5 µg/ml Blasticidin. For splitting, cells were washed with PBS and incubated with trypsin for 5 - 10 min at 37 °C before dissolving the cells in fresh media and plating an appropriate number of cells in a new plate.

In order to adapt adherent cells to suspension growth about $7*10^7$ cells were detached from plates with trypsin as for splitting, but then resuspended in 100 ml DMEM high glucose media supplemented with 1 % FBS and 1 % Penicillin/Streptomycin and cultivated in spinner flasks under moderate stirring (25 - 40 rpm) at 37 °C and 5 % CO_2 atmosphere. To grow a larger amount of cell mass, the culture was expanded with 1x the current volume once the cells reached a density of $\sim 7*10^5$ cells/ml and transferred to a new spinner flask of increasing volume when required.

Cells for protein purification were harvested in a small scale from adherently grown cells and for larger scale from suspension culture. Adherently grown cells were manually detached with a cell scraper and resuspended in media first. Afterwards adherent cells and suspension culture cells were treated identically. Cells were pelleted by centrifugation at 2 500 g and 4 °C, washed with PBS. The pellets were flash-frozen in liquid nitrogen and stored at -80 °C.

6.5.2 Large-scale fermentation

With the help of the Archaea Centre Regensburg, human suspension cell culture was upscaled. About $12*10^9$ cells were inoculated with 70 L DMEM high glucose media supplemented with 1 % FBS and 1 % Penicillin/Streptomycin and incubated at 37 °C and 5 % CO_2 atmosphere at 40 rpm and 1050 ml/min aeration until a cell density of $1*10^6$ cells/ml was reached and cells were harvested by centrifugation. Cell pellets were flash-frozen in liquid nitrogen and stored at -80 °C.

6.5.3 Purification of genomic DNA

Cell pellets (about $1*10^6$ cells) were resuspended in 400 µl Proteinase K buffer and incubated with 0.2 mg/ml Proteinase K overnight at 50 °C. Afterwards isopropanol precipitation was performed before resuspending the DNA with MilliQ water. The human genomic DNA was stored at 4 °C.

6.5.4 Denaturing protein extraction

Adherently growing cells (about $3*10^6$ cells) were detached with 300 µl boiling 1x SDS loading dye and rigorously shaken for 15 min at 95 °C.

6.5.5 RNA extraction from human cells

RNA extraction was performed from HeLa or HEK cells with the NucleoSpin RNA clean-up kit (Macherey-Nagel) according to manufacturer's instructions.

6.5.6 Transfection

Adherently grown cells were transfected with plasmid DNA with the help of FuGENE HD Transfection Reagent according to manufacturer's instructions.

6.5.7 Genome editing by CRISPR/Cas9

Genome editing by CRISPR/Cas9 was performed as described[179,224], similar to a previously published protocol with some modifications[245]. In general, a Cas9 nickase (Cas9n) was used with two different guide RNAs (gRNAs) to generate a staggered double-strand break with single-stranded overhangs on the C-terminus of the desired gene. This approach minimizes off-target effects, because at all potential off-target sites of one gRNA only a single-strand break is introduced in the genome. Only at the specific desired site Cas9n cuts both strands in close proximity due to the two different gRNAs. Once the cell recognizes a double-strand break within the genome, two repair mechanisms are possible: 'non-

homologous end joining' (NHEJ) and 'homology-directed repair' (HDR). NHEJ is a random repair mechanism during which insertions or deletions of a few base pairs (bp) at the double-strand break are common. In contrast, the HDR can be used for insertion of a specific DNA sequence. Here, a DNA template, containing homologous DNA of both sites of the to-be-introduced sequence, is used to repair the double-strand break and results in an exact copy of the template DNA. This method was used to tag hPol I subunit RPA1 (gene *POLR1A*) endogenously in a cell line. For this purpose, template DNA, containing two homology arms flanking the sequence which should be inserted was transfected into the cells during genome editing as a plasmid. The homology arms consist of 800 bp each, one containing the coding sequence ending with the last codon, just before the stop codon and the other one starting with the stop codon going further in the 3' untranslated region (3' UTR). The introduced sequence was composed of a GS-linker with a 3C protease cleavage site followed by the open reading frame (ORF) of the superfolder GFP (sfGFP). At least one of the two PAM sequences of the gRNAs were silently mutated within the homology arms to avoid continuing cutting of Cas9n after successful introduction of the desired sequence.

For gRNA design the web-based tool Benchling was used to select an appropriate pair of gRNAs. Annealed oligonucleotides containing the gRNA sequence were cloned via BbsI as described in the manual into the Cas9n expression vector pSpCas9n(BB)-2A-Puro (pX462) V2.0. To avoid the necessity of transfecting two plasmids with Cas9n and crRNA (CRISPR-RNA), one crRNA was subcloned by PCR, restriction digestion and ligation together with its promoter and terminator in the second pX462 plasmid. As a second plasmid the donor plasmid was cloned to allow tag insertion by HDR of the cell. First, a 1.6 kb long DNA piece with the stop codon of the desired gene right in the middle was cloned from genomic DNA. Just after the codon of the last amino acid and before the stop codon the sequence of a short GS-linker with an embedded HRV 3C protease cleavage site and the sfGFP ORF was introduced. Additionally, the PAM sequence of at least one gRNA was altered with a silent mutation in order to avoid further cutting by Cas9 after the tag was successfully inserted.

Adherent cells were transfected with equimolar amounts of plasmids as described in 6.5.6. Several days after transfection the GFP-expressing cells were enriched by fluorescence activated cell sorting (FACS) using a BD FACSAria™ IIu cell sorter at the Central FACS Facility of the RCI Regensburg (Center for Interventional Immunology). GFP-positive cells were seeded as single cells on 96-well plates. After 2 - 3 weeks, colonies were expanded. These monoclonal populations were validated for homozygous tag insertion using PCR and subsequent sequencing and Western blot. For this purpose, genomic DNA was extracted and a PCR spanning the C-terminus of the gene of interest performed. The PCR product was then run on an agarose gel, bands were gel extracted and sequenced. For Western blotting, denaturing protein extraction of the cells were performed, ran on a SDS gel, transferred on a membrane by Western blotting and the protein of interest directly detected by specific antibodies. The tagged protein was enlarged by the size of the tag compared to the parental cell line.

6.5.8 Establishing stable overexpression cell lines with the T-REx system

Stable, inducible overexpression cell lines were generated with the T-REx system (Invitrogen) according to manufacturer's instructions. For this purpose, the open reading frame (ORF) of the protein of interest was cloned with a C-terminal GS-linker containing a 3C cleavage site followed by a sfGFP ORF as well as a FLAG-HA tag into the pcDNA5/TO plasmid. This plasmid was transfected in a 9:1 molar ratio of pOG44:pcDNA5/TO plasmid into T-REx-293 cells. The following day, the media was replaced by fresh media. The day after, cells were split to 25 % confluency at a maximum and Hygromycin B (100 µg/ml final concentration) and Blasticidin (5 µg/ml final concentration) was added to the media for selection. Every 3 - 4 days media was exchanged to fresh media additionally containing Hygromycin B and Blasticidin. Once single colonies have formed, they were separated from each other and expanded separately. Successful integration for each clone was verified by induction with tetracycline, which led to protein induction.

For overexpression, induction with tetracycline (1 µg/ml end concentration) for two days was performed to allow the other subunits of the protein complex to be produced by the cell as well. After two days cells were harvested by centrifugation and cell pellets flash-frozen in liquid nitrogen and stored at -80 °C.

6.5.9 Cell cycle synchronization

Synchronization of cell cycle was performed as previously described[256]. Cells were grown adherently to nearly 50 % confluency, before cell culture medium was supplemented with 2 mM Thymidine for 16 h, to arrest the cell cycle. Afterwards cells were washed four times with medium and further grown. About ten hours later roughly synchronized cells started mitosis.

6.5.10 Cell fixation and preparation for immunofluorescence

Preparation of cells for immunofluorescence (IF) was essentially performed as described[179,224]. Cells were grown adherently on glass cover slips to 50 % confluency and after washing the cells with pre-warmed (37 °C) PBS, they were fixed with 3.7 % paraformaldehyde in PBS for 10 min at 37 °C. The fixation was stopped by replacing the solution with 100 mM glycine in PBS for 5 min at 37 °C. Afterwards, cells were washed twice with PBS, mounted on the specimen slide with the help of a drop of Prolong Gold Antifade Mountant with DAPI, and dried in the dark at least overnight.

If additional epitopes were detected by antibodies, additional staining steps were performed after the final washing step, just before mounting. First, the cells were incubated with 0.2 % Triton-X in PBS for 15 min at room temperature for permeabilization and washed three times with IF wash buffer. For staining, incubation in IF wash buffer with primary antibody diluted as listed was performed for 1 h. Subsequently, the cells were washed four times with IF wash buffer before they were incubated with IF wash buffer containing accordingly diluted secondary antibody for 1 h. Finally, the cells were washed with IF wash buffer and PBS. In both cases, the cells on the cover slips were finally washed with water before mounting.

6.6 Purification of hPol III

Purification of hPol III was essentially performed as described[179] and similar to a published protocol[224]. The HeLa-POLR1C cell pellet was resuspended in hPol Lysis buffer supplemented with 7 U/ml DNase I and lysed by Dounce homogenization and incubation on ice for 30 min. After centrifugation at 20 000 g for 15 min at 4 °C, the lysate was incubated for 3 h at 4 °C with pre-equilibrated GFP-Trap Dynabeads under continuous rotation (7 rpm). Beads were washed twice with hPol Wash buffer and eluted with 1x slurry volume hPol Wash buffer supplemented with 10 µg of 3C protease per 1 g of cell pellet for 4 h at 4 °C. The eluate was diluted with hPol Mono A buffer to 140 mM NaCl before loading on a MonoQ 1.6/5 PC column equilibrated with 60 mM ammonium sulfate. Elution was performed with a linear gradient over five column volumes to 200 mM followed by steps of five column volumes with 200 mM, 350 mM, 600 mM and 1 M ammonium sulfate. hPol III eluted at 600 mM ammonium sulfate concentration (hPol I peak at 350 mM ammonium sulfate) and was immediately used for grid preparation or in transcription assays.

6.7 Purification of hPol I

hPol I purification was performed as described[224]. HeLa-POLR1A cell pellet was resuspended in twice the volume of the cell pellet's weight of hPol Lysis buffer supplemented with 7 U/ml DNase I and lysed by Dounce homogenization and incubation on ice for 30 min. After centrifugation at 20 000 g and 4 °C for 15 min, the whole-cell lysate was incubated with pre-equilibrated GFP-Trap Dynabeads for 3 h. The beads were washed once with four times and once with twice the slurry volume of hPol Wash buffer, before being eluted with the slurry volume of hPol Wash buffer supplemented with 10 µg of 3C protease per 1 g of cell pellet for 4 h at 4 °C.

In case an anion-exchange chromatography was performed, the GFP-elution was diluted with hPol Mono A buffer to reach a final concentration of 140 mM NaCl. The sample was loaded on a MonoQ 1.6/5 PC column equilibrated with 60 mM ammonium sulfate and eluted stepwise in buffer A with increasing the concentration of ammonium sulfate up to 1 M. A linear gradient over five column volumes to 200 mM followed by steps of five column volumes with 200 mM, 350 mM, 600 mM and 1 M ammonium sulfate was applied. hPol I eluted at 350 mM ammonium sulfate concentration. hPol I was used immediately or flash-frozen in liquid nitrogen and stored at -80 °C for further experiments.

6.8 Purification of Pol I subunits and domains

6.8.1 Purification of human RPA49/34

Purification of human RPA49/34 subunits was performed as described[224]. The different variants of the human heterodimer ($RPA49^{FL}/34^{FL}$, $RPA49^{FL}/34^{1-343}$, $RPA34^{131-510}$) were cloned with a N-terminal 6xHis-tag on RPA49 and untagged RPA34, except for $RPA34^{131-510}$, which carries an N-terminal 6xHis-tag itself. The proteins were co-expressed in *E. coli* BL21 (DE3) RIL in LB medium with 0.2 mM IPTG overnight at 18 °C. Cells were resuspended in RPA49/34 Lysis buffer and lysed by sonification. After centrifugation,

the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed subsequently with six times the bed volume of RPA49/34 Wash I buffer, RPA49/34 ATP-Wash buffer, another ATP-Wash after 10 min of incubation and RPA49/34 Wash II buffer before being eluted with RPA49/34 Elution buffer. The ATP-Wash steps were performed at room temperature. The sample was diluted 5-fold with RPA49/34 Mono A buffer before loading onto a MonoS 5/50 GL column with Mono A buffer supplemented with 100 mM NaCl. Elution was performed with a linear gradient of NaCl concentration up to 2 M. The corresponding fractions were pooled and concentrated with 10 kDa cut off and applied to a Superdex200 Increase 10/300 column equilibrated with RPA49/34 SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

6.8.2 Purification of yeast A49/34.5

The *S. cerevisiae* A49/34.5 was purified as described[56,224]. Full-length heterodimer A49/34.5 was co-expressed with a C-terminal 6xHis-tag on Sc A49 and untagged Sc A34.5 in *E. coli* BL21 (DE3) RIL in LB medium with 0.2 mM IPTG for 18 h at 18 °C. The cells were resuspended in A49/34.5 Lysis buffer and sonified. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed with six times the bed volume of A49/34.5 Wash I buffer and of A49/34.5 Wash II buffer before being eluted with A49/34.5 Elution buffer. The sample was diluted 3-fold with A49/34.5 Dilution buffer before loading onto a MonoS 5/50 GL column with A49/34.5 Mono A buffer. Elution was performed with a linear gradient of NaCl concentration up to 1 M. Sc A49/34.5 eluted at around 280 mM NaCl. The corresponding fractions were pooled and concentrated with 10 kDa cut off and applied to a Superdex200 Increase 10/300 column equilibrated with A49/34.5 Mono A buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

6.8.3 Purification of human RPA12

Human RPA12 was cloned with a N-terminal 6xHis-tag and expressed in *E. coli* pRARE cells in LB medium. When cells were grown to on OD_{600} of 1.0, the culture was cooled down for 15 min in an ice bath and induced for 3 h at 18 °C with 0.1 mM IPTG. Cells were resuspended in RPA12 Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed with six times the bed volume of RPA12 Lysis buffer and RPA12 Wash buffer before being eluted with RPA12 Elution buffer. The eluate was diluted to 150 mM KCl with RPA12 Mono A buffer and loaded on a HiTrap Q column. Elution was performed with a linear gradient of KAc concentration up to 2 M. RPA12 elutes at 450 mM KAc and corresponding fractions were pooled and concentrated with 3 kDa cut off and applied to a Superdex75 Increase 10/300 GL column equilibrated with RPA12 SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80°C.

6.8.4 Purification of yeast A12.2

Purification of yeast subunit A12.2 was essentially performed as described[51]. *S. cerevisiae* A12.2 was cloned with a N-terminal 6xHis-tag and expressed in *E. coli* BL21 (DE3) pRARE in LB medium. When cells were grown to on OD_{600} of 1.0, the culture was cooled down for 15 min in an ice bath and induced for 3 h at 18 °C with 0.1 mM IPTG. Cells were resuspended in RPA12 Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed subsequently with six times the bed volume of RPA12 Lysis buffer, RPA12 Wash buffer,

RPA12 ATP-Wash buffer, another ATP-Wash after 10 min of incubation and RPA12 Wash buffer before being eluted with RPA12 Elution buffer. The ATP-Wash steps were performed at room temperature. The eluate was diluted to 150 mM KCl with RPA12 Mono A buffer and loaded on a HiTrap Heparin column. Elution was performed with a linear gradient of KAc concentration up to 2 M. A12.2 is not binding to the column under these conditions and flow-through fractions were pooled and loaded on a HiTrap Q column. A12.2 was found in the flow-through fractions and these were pooled and concentrated with 3 kDa cut off and applied to a Superdex75 Increase 10/300 GL column equilibrated with RPA12 SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

6.8.5 Purification of human dock II domain of RPA1

Purification of human dock II domain was performed as described[224]. Two variants of the human dock II domain ($RPA1^{1060-1155}$ (full-length), $RPA1^{1081-1146}$ (minimal)) were cloned with a C-terminal 6xHis-MBP-tag (3C cleavable) and expressed overnight at 20 °C in *E. coli* BL21 (DE3) RIL in LB medium with 0.2 mM IPTG as well as 6xHis-MBP-tag-only. Cells were resuspended in Dock II Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed subsequently with six times the bed volume of Dock II Wash I buffer and Dock II Wash II buffer before being eluted with Dock II Elution buffer. The eluate was buffer-exchanged to Dock II SEC buffer with a PD10 column and applied to a Superdex 75 Increase 10/300 GL column equilibrated with SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

6.9 Purification of hPol I transcription initiation factors

6.9.1 Purification of human RRN3

Human RRN3 was cloned for two different expression and purification approaches. A 3C-cleavable N-terminal 6xHis-MBP-tag was cloned for expression in SF21 insect cells and a N-terminal 6xHis-tag for bacterial expression.

Expression of human RRN3 was performed in insect cells SF21 as described (see method section 6.4) Cells were resuspended in Rrn3-SF Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded on an MBPTrap column and eluted with Rrn3-SF Elution buffer. The eluate was concentrated with a 30 kDa cut off filter and applied to a Superdex200 Increase 3.2/300 column equilibrated with Rrn3-SF SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

Expression in bacterial cells and purification of human RRN3 was essentially performed as described[222]. For bacterial expression *E. coli* BL21 (DE3) R3 pRARE cells were grown in TB medium at 18 °C until they reached on OD_{600} of 0.8 - 1.0 and induced with 0.05 mM IPTG for 16 h at 18 °C. Cells were resuspended in Rrn3 Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed subsequently with six times the bed volume of Rrn3 Lysis buffer, Rrn3 Wash I buffer, Rrn3 Wash II buffer, Rrn3 Wash III buffer and after 10 min of incubation Rrn3 Wash IV buffer before being eluted with Rrn3 Elution buffer. The eluate was loaded onto a HiTrap Q column with Rrn3 Mono A buffer supplemented with 200 mM NaCl. Rrn3 was bound

to the column and eluted with increasing amounts of NaCl up to 1 M. RRN3-containing fractions were pooled and concentrated with a 30 kDa cut off filter and applied to a Superdex200 Increase 10/300 column equilibrated with Rrn3 SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80°C. Further optimization of this protocol in future might be performed in order to enhance purity of the sample.

6.9.2 Purification of human SL1

Human SL1 purification was performed from cell pellets of T-REx-293-TAF1A cells in which overexpression had been induced for two days. The cell pellet was resuspended in twice the volume of the cell pellet's weight of SL1 Lysis buffer supplemented with 50 U/ml DNase I and incubated on ice for 15 - 30 min. After sonification, the cells were incubated for another 15 - 30 min on ice. After centrifugation at 20 000 g and 4 °C for 15 min, the whole-cell lysate was incubated with pre-equilibrated GFP-Trap Dynabeads for 3 h. The beads were washed once with four times and once with twice the slurry volume of SL1 Wash buffer, before being eluted with one slurry volume of SL1 Wash buffer supplemented with 30 µg of 3C protease per 1 g of cell pellet for 4 h at 4 °C. The SL1-containing eluate was incubated with pre-equilibrated MagneHis Ni-Particles for 1 h. SL1 is found in the supernatant, whereas the 3C protease is bound to the beads.

6.9.3 Purification of human UBF1 and UBF2

UBF1 and UBF2 were cloned with a N-terminal 6xHis-tag and expressed in *E. coli* BL21 (DE3) R3 pRARE in LB medium. Freshly washed cells were grown at 30 °C until OD_{600} reached 1.0 and expression was induced with 0.25 mM IPTG for 1 h at 30 °C. Cells were resuspended in UBF Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed subsequently with six times the bed volume of UBF Lysis buffer, UBF Wash I buffer, UBF Wash II buffer, UBF ATP-Wash buffer and another ATP-Wash after 10 min of incubation before being eluted with UBF Elution buffer. The ATP-Wash steps were performed at room temperature. The sample was diluted 2-fold with UBF Mono A buffer before loading onto a HiTrap Heparin column with UBF Mono A buffer supplemented with 150 mM KAc. Elution was performed with a linear gradient of KAc concentration up to 2 M. The corresponding fractions were pooled and concentrated with a 30 kDa cut off filter and applied to a Superdex200 Increase 10/300 column equilibrated with UBF SEC buffer. Pooled peak fractions were concentrated and flash-frozen for storage at -80 °C.

6.10 Purification of human Top2a domains

Expression plasmids for the purification of human Top2a ATPase domain (aa 29 - 428) and catalytic domain (aa 431 - 1217) were obtained from Prof. Dr. Valerie Lamour. Top2a-ATPase was cloned with a N-terminal 6xHis-MBP-tag and Top2a-cat with a 3C cleavable 6xHis-tag.

Top2-ATPase expression and purification were essentially performed as published[257]. It was expressed in *E. coli* BL21 (DE3) RIL in LB medium. At an OD_{600} of 0.8, protein expression was induced by 0.4 mM IPTG for 10 h at 30 °C. Cells were resuspended in ATPase Lysis buffer and lysed by sonification. After

centrifugation, the lysate was loaded on a HisTrap column, washed subsequently with five column volumes of ATPase Lysis buffer and ATPase Wash buffer and eluted with ATPase His-Elution buffer. The eluate was loaded on a HiTrap Heparin column with ATPase Mono A buffer supplemented with 150 mM NaCl. Elution was performed with a single step of 450 mM NaCl concentration. The corresponding fractions were pooled and loaded on a MBPTrap column with ATPase Mono A buffer supplemented with 450 mM NaCl. Top2a-ATPase was eluted with ATPase MBP-Elution buffer and concentrated with a 30 kDa cut off filter while buffer exchanged to ATPase MonoA buffer supplemented with 450 mM NaCl and flash-frozen for storage at -80 °C.

Top2a-cat expression and purification protocol were based on the purification protocol of the catalytic domain of Top2b[258]. Top2a-cat was expressed in *E. coli* BL21 (DE3) RIL in LB medium. At an OD_{600} of 0.6, protein expression was induced with 0.3 mM IPTG for 16 h at 20 °C. Cells were resuspended in Top2a-cat Lysis buffer and lysed by sonification. After centrifugation, the lysate was loaded onto pre-equilibrated Ni-NTA beads by gravity-flow, washed with six times the bed volume of Top2a-cat Lysis buffer and eluted with Top2a-cat Elution buffer. The eluate was dialyzed against Top2a-cat Mono A buffer overnight and loaded on a HiTrap Heparin column with Top2a-cat Mono A buffer. Elution was performed with a linear gradient to a NaCl concentration of up to 1 M. The corresponding fractions were pooled and concentrated with a 100 kDa cut off filter and applied to a Superose6 Increase 10/300 equilibrated with Top2a-cat SEC buffer. Pooled peak fractions were concentrated and flash-frozen in liquid nitrogen for storage at -80 °C.

6.11 Electron microscopy

6.11.1 Protein batch crosslinking

Protein batch crosslinking was performed with glutaraldehyde directly before grid preparation. To a 0.1 - 0.3 mg/ml protein sample, 0.1 % glutaraldehyde end concentration was added and incubated for 3 - 7 min before applying the sample directly on the grid without stopping the crosslinking reaction.

6.11.2 Negative stain grid preparation

Negative stain grids were prepared as described[224]. Protein samples were centrifuged before 5 µl were applied on glow-discharged 400-mesh copper grids (G2400C; Plano) with a self-made carbon film of ~ 7 nm thickness[188]. After 30 s - 1 min, grids were washed with ddH_2O for 30 s, and stained three times with 5 µl saturated uranyl formiate solution (2x 20 s, 1x 30 s). After each step, excess liquid was removed with a filter paper.

6.11.3 Preparation of carbon-supported grids

Grids for cryo-EM were prepared with a self-made carbon supported film as published[188]. Thin carbon support films of ~ 2 nm were made with the Cressington Turbo Carbon Coater 208carbon on top of a fresh fissure of a mica crystal. Glow-discharged grids were submerged on a filter paper in ddH_2O and the carbon support film was floated on the water surface. By slowly draining the water chamber, the film was placed on the grids, which were dried at least overnight at room temperature before usage.

6.11.4 Preparation of graphene oxide (GO) supported grids

Different protocols were tested for self-made graphene-oxide (GO) supported grids[259–261], but only one resulted in reproducible grid quality[261]. Grids were washed twice with a drop of chloroform before glow-discharging them for 30 s. 4 µl of polyethylenimine (PEI, 1 mg/ml in 25 mM Hepes pH 7.0) was applied to the grid and incubated for 2 min. After washing the grid twice with 4 µl of ddH_2O and blotting away excess liquid with filter paper in between, grids were dried for 2 min on filter paper before GO coating. For this purpose, GO solution was diluted to 0.2 mg/ml with water and centrifuged for 1 min at 1 500 g. Aggregated GO in the pellet was discarded and 4 µl of the supernatant was applied to the grid for 2 min. Afterwards, grids were washed twice with 4 µl of ddH_2O and excess liquid was blotted away with filter paper in between. Grids were dried for 2 min on a filter paper and used for cryo-EM grid preparation within the next 3 h.

6.11.5 Electron cryo-microscopy (cryo-EM) grid preparation

Preparation of cryo-EM grids was essentially performed as described[224]. 0.1 - 0.3 mg/ml protein sample was used for supported cryo-EM grids. 3 µl of sample were applied to the grids and incubated for 30 s - 1 min at 100 % humidity and 4 °C in a Vitrobot mark IV and plunged into liquid ethane after blotting. Blotting was depended on support film, GO-supported grids were blotted for 2.5 s with blot force 7, whereas carbon-supported grids were blotted for 2.5 s with blot force 9. For unsupported cryo-EM grids, the protein sample (up to 1 mg/ml concentration) was not incubated on the grid and blotted for 5 s with blot force 12. Prepared cryo-EM grids were always handled within liquid nitrogen and stored in liquid nitrogen until usage.

6.11.6 Electron microscopy of negative stained samples and processing

Electron microscopy of negative stained grids was performed as described[224]. The images were collected on a JEOL 2100-F Transmission Electron Microscope operated at 200 keV and equipped with TVIPS-F416 (4k x 4k) CMOS-detector at 40 000 x magnification (pixel size 2.7 Å) with alternating defocus (-1 to -3 µm). The images were processed using RELION 3.1[190]. For hPol I samples, a total of 76 micrographs were analyzed, yielding 46 196 auto-picked particles using the Laplacian-of-Gaussian (LoG) routine followed by reference-free 2D sorting, 3D classification (reference PDB: 5M3M low-pass filtered to 60 Å) and 3D refinement.

6.11.7 Electron cryo-microscopy data collection and processing of hPol I

Cryo-EM data collection and processing of hPol I was performed as described[224]. A total of 9 709 micrograph movies from GO-supported cryo-EM grids of hPol I protein sample were collected on a CryoArm200 cryo-electron microscope (JEOL) equipped with a K2 direct electron detector (Gatan), in-column energy filter and Cold-Field Emission Gun (low-flash interval 4 h). A total dose of 40 $e^-/ Å^2$ was fractionated over 40 frames at a defocus range of -1.2 to -2.7 µm using SerialEM[262] in a 5x5 multi-hole strategy as described[263].

Pre-processing was carried out using WARP[189], followed by 2D and 3D classification and auto-refinement using RELION 4.0[190]. During pre-processing motion-correction, CTF estimation and particle

picking was performed. The pixel size was binned to 1.50846 Å/pix and particles extracted with a box size of 190 pixel. Rough 2D classification followed by 3D classification using a reference of hPol I obtained after stringent 2D classification and 3D refinement yielded a reconstruction at an overall resolution of 4.09 Å. Further 3D classification was performed to investigate the occupancy and flexibility of the dimerization domain of RPA49/34 and the clamp/stalk region.

6.11.8 Model building and refinement of hPol I

Model building and refinement of hPol I was performed as described[224]. Models for common subunits RPABC1-5 and the RPAC1/2 assembly were transferred from an hPol III reconstruction[191]. Homology models of the hPol I subunits RPA1, RPA2, RPA49, RPA34, RPA12 and RPA43 were generated based on sequence and secondary structure alignments with the crystal structures of their *S. cerevisiae* counterparts (see result section 7.1) using the MODELLER software package[192]. The models were adjusted in COOT[264] and real-space refined using Phenix[193]. At later stages, released AlphaFold[194] models were used to guide chain-tracing in poorly resolved areas and specifically modelling of the dock II domain was supported by its AlphaFold prediction. The model of stalk subunit RPA43 is solely based on structure predictions and is included in some figures but was not deposited due to poor or absent cryo-EM density resulting from flexibility.

6.12 Functional and biochemical assays

6.12.1 Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays (EMSAs) were performed as described[224] in order to visualize *in vitro* binding of a purified protein to DNA and estimate the binding affinity. A total of 100 fmol fluorescently labelled, double-stranded DNA were mixed with up to 50-fold molar excess of purified protein (as labeled in the figure) in EMSA buffer. RPA49/34 DNA-binding was performed with a pre-annealed, 40 bp long, randomized DNA and EMSA-NaCl and EMSA-NaCl-2 buffer. A 339 bp long *S. cerevisiae* rDNA promoter sequence was used to test UBF binding in EMSA-NaCl, EMSA-KCl and EMSA-$(NH_4)_2SO_4$ buffer. After incubation at room temperature for 30 min, 6x NP OrangeG loading dye was added to reach 1x concentration. 5 - 10 % native polyacrylamide gels in 0.4x TBE were pre-run at 110 V for 30 min before the reaction was separated at 110 V for 1:45 h at 4°C. The Cy5-labeled DNA was detected with a Typhoon FLA9500.

6.12.2 *In vitro* transcription assay

In vitro RNA elongation and cleavage assays were essentially performed as described[179,224]. 0.5 pmol of Pol I from *S. cerevisiae*, *S. pombe*, or *H. sapiens* were pre-incubated with 0.25 pmol of different pre-annealed minimal or bubble nucleic acid scaffolds (sequence information schematically shown in each figure along with the gel) in Elongation/Cleavage buffer for 1 h at 20 °C in a total volume of 45 µl. In case purified RPA49/RPA34 heterodimer was added, 1x, 5x or 10x molar excess of heterodimer compared to polymerase was included during the pre-incubation. For RNA elongation, 10 µmol of each NTP (which NTP(s) is specifically marked at each lane in the corresponding figure) were added and the

reaction was incubated for 1 h at 28 °C. To examine cleavage activity, the pre-incubated reaction was incubated for 1 h at 28 °C without the addition of NTPs. Afterwards, nucleic acid purification was examined by adding 5 M NaCl to a final concentration of 0.5 M and 800 µl 100 % ethanol. After precipitation for at least 1 h at -20 °C, the sample was centrifuged for 30 min at 20 000 g and 4 °C. The pellet was washed with 80 % (v/v) ethanol and, after drying, resuspended in 1x RNA loading dye (xylene cyanol only for FAM-labeled constructs included in buffer). The sample was heated to 95°C for 5 min. As control 0.25 pmol of scaffold were treated identically, without addition of polymerase and NTPs. 0.125 pmol of RNA product were separated by gel electrophoresis and visualized with a Typhoon FLA9500.

6.12.3 Tail template transcription assay

Tail template transcription assays were essentially performed as described[64]. A 2 kb long PCR product (template DNA: plasmid 2318; Primer: 4220 and 4019) was digested with the nickase Nb.BsmI. After denaturation of the enzyme, competitor Oligonucleotide (Primer 2207) was added in 10x molar excess, and the sample heated for 10 min at 80 °C before cooling down slowly to room temperature. The tail template DNA was precipitated with ethanol as described in 6.1.11 and dissolved in water to reach a concentration of 250 nM. For one reaction, 0.25 pmol Pol I, 50 pmol GpC and 0.125 pmol tail template DNA were mixed in 50 mM Hepes pH 7.8. The reaction was started with TT transcription buffer (1x final concentration) supplemented with 0.3 µl α-^{32}P-CTP for 15 min at 24 °C. During this time, Pol I transcription was started until a specific stalling site. Pol I was paused at this position as no GTP was in the reaction which would be needed at this site repetitively. This approach allows to start Pol I transcription elongation at the same time from a specific site for all enzymes by adding TT Start buffer to the reaction to reach 1x concentration. At different time points, the reaction was stopped by adding RNA loading dye (tail template) to reach 1x concentration. Samples were heated for 5 min at 95 °C before loading on a tail template Urea-gel. The gel was pre-run for 20 min at 25 W and the RNA samples were separated for 1 h at 25 W in 1x TBE running buffer. Afterwards the gel was dried with a Model 583 Gel Dryer at 80 °C for 2 h under vacuum and the signals were detected with the help of an Imaging Screen-K and the PMI Personal Molecular Imager FX Scanner.

6.12.4 Dock II-Top2a pulldowns

To investigate direct protein-protein interactions between the dock II domain and Top2a, pulldown assays with immobilized dock II domain were performed. 100 µg of dock II domain versions (MBP-only, MBP-dock II, MBP-minimal dock II) were bound to 10 µl magnetic His-beads in 300 µl Dock II-Top2a binding buffer for 45 min at 4 °C and 7 rpm. Afterwards, excess protein was washed away three times with 200 µl Dock II-Top2a binding buffer. Now, 3 µg of full-length Top2a was added to the coupled beads in 100 µl Dock II-Top2a binding buffer for 3 h at 4 °C and 7 rpm. After three washing steps with 50 µl Dock II-Top2a binding buffer, proteins were eluted with imidazole from the beads with 20 µl of Dock II-Top2a elution buffer. Afterwards, another elution step under high-salt conditions was performed with 50 µl Dock II-Top2a high salt buffer, before beads were resolved in 20 µl 1x SDS loading dye to analyze remaining proteins. SDS-PAGE electrophoresis and Coomassie-staining of the different

samples as well as input samples was executed, to analyze which proteins elute in which step and if direct protein-protein interaction is possible.
For variations MBP-beads to couple Dock II variants (elution via maltose or 3C protease cleavage), StrepTactin beads to couple Top2a, or addition of BSA to unspecifically saturate beads were tried.

6.12.5 Blue native PAGE

To investigate protein-protein interaction, blue-native PAGE was performed as described[224]. Five times molar excess of MBP-tag only or tagged human dock II domain was incubated with recombinant Top2a ΔC (1-1217, a gift from Prof. Valerie Lamour) in Protein-Protein binding buffer for 30 min at room temperature. After adding NativePAGE sample buffer, the samples were separated on a Native PAGE 3-12 % gradient gel at 150 V for 90 min with light blue cathode and anode buffer according to the manufacturer's instructions and Coomassie stained for protein detection.

6.12.6 Mass photometer

The Mass photometer was used together with the AcquireMP and DiscoverMP software as recommended by the manufacturer (Refeyn). Protein samples were diluted to 100 - 500 nM and the device was size calibrated with BSA, IgG and Thyroglobulin. To find the focus a droplet of the corresponding protein buffer was pipetted in one well, and after adding the protein sample to an end concentration of 10 - 40 nM, the mass of the protein was measured for 1 min.

6.13 Confocal microscopy

Confocal microscopy was performed as described[224] with support and help from Dr. Andrea Bleckmann. The fluorescent specimens were imaged using a Plan-Apochromat 63x/1,4 Oil DIC Objective at a Zeiss LSM980/Airyscan 2 confocal microscope. sfGFP was excited by a 488 nm diode laser and emission was detected using a 300 - 720 nm band pass filter. Separately, DAPI was excited by a 405 nm diode laser and emission was detected using a 300 - 720 nm band pass filter. For the 3D model a Z-stack was imaged using the internal GaAsP-PMT detectors from 490 - 668 nm for sfGFP and 410 - 473 nm for DAPI in a two-track process. Image processing was done using the Zeiss AxioVision and Fiji software. The 3D Volume images were created in Imaris 9.6.

6.14 Experiments performed in cooperation

6.14.1 Phylogenetic analysis

Phylogenetic analysis was performed by Kristina Straub as described[224].

Analysis of Pol I subunits RPA1, RPA34, RPA43 and A14

Data sets from Pol I subunits were generated using their corresponding InterPro[265] entries (RPA1: IPR015699, RPA34: IPR013240, RPA43: IPR041901 and IPR041178, A14: IPR013239 downloaded on 07.06.2021). A common dataset of RPA1, RPA34 and RPA43 was generated by searching for common species within the three InterPro families. To each obtained species the concatenated sequence of RPA1, RPA34 and RPA43 was assigned.

Phylogenetic analysis

Sequence alignment tool MAFFT[266] has been used with default options and a gap open penalty of 70. The resulting alignment was filtered manually on highly diverged sequences. To improve the quality of the phylogenetic analysis without losing information for each genus only one sequence was chosen. On the resulting data set with 513 sequences Gblocks[267] (options: b3=5000, b4=2, b5=a) has been applied to remove uninformative columns. By means of RAxML[268] using the option '-f a' and the substitution model PROTGAMMAAUTO 100 trees were generated and a consensus tree was derived. The root has been placed between the supergroups of *Sar* and *Haptophyta* and the supergroup of *Amorphea*[269]. The resulting phylogenetic tree was analyzed with respect to the taxonomic distribution. Sequences were grouped according to branching points in the phylogenetic tree (Figure 11). In order to retrieve the taxonomic group where the A14 subunit is present the species related to the A14 subunit InterPro entry are compared with the species given in the phylogenetic tree.

Sequence analysis of RPA34 and RPA1

By means of MAFFT sequence alignment of each subunit was generated using varied gap open penalties (RPA34: 50, RPA1: 20). Due to higher sequence variety within RPA34 sequences BLOSUM30 was used instead of the default parameter. In order to account the divergence between the taxonomic groups given from the phylogenetic tree, the alignment was split into these groups and each group was analyzed separately on the presence or absence of the RPA34 C-terminal extension, the RPA1 foot domain and the RPA1 expander domain. Sequences from Homo sapiens have been used as reference to identify the region of interests (RPA34 C-terminal extension: 399-510; RPA1 foot domain: 1074-1139; RPA1 expander domain: 1365-1488). The median length and standard deviation of the regions of interest have been calculated for each group. To unravel the sequence and structural conservation of the regions of interest the conservation score given in Jalview[270] has been extracted after removing all columns containing only gaps. The mean conservation score is calculated by summing up over all column scores divided by the number of columns. Scores are grouped into 5 categories: not conserved (0-3), weakly conserved (3-5), medium conserved (5-7), conserved (7-9), strongly conserved (9-11). Secondary structures were predicted using Ali2D[195,271]. Secondary structure elements were assigned when more than 5 amino acids have medium to high probability in more than 90 % of the sequences

within each group. Bridging of two secondary structure elements over less than 5 differently annotated amino acids are counted as one element. If gaps are present in more than 90 % of the sequences, they are ignored.

6.14.2 Mass spectrometry

Mass spectrometry was performed by the group of Dr. Astrid Bruckmann as described[224]. Protein bands were cut out from the gel, washed with 50 mM NH_4HCO_3, 50 mM NH_4HCO_3/acetonitrile (3/1), 50 mM NH_4HCO_3/acetonitrile (1/1) and lyophilized. After a reduction/alkylation treatment and additional washing steps, proteins were in-gel digested with trypsin (Trypsin Gold, mass spectrometry grade, Promega) overnight at 37 °C. The resulting peptides were sequentially extracted with 50 mM NH_4HCO_3 and 50 mM NH_4HCO_3 in 50 % acetonitrile. After lyophilization, peptides were reconstituted in 20 µl 1 % TFA and separated by reversed-phase chromatography. An UltiMate 3000 RSLCnano System (Thermo Fisher Scientific) equipped with a C18 Acclaim Pepmap100 preconcentration column (100 µm i.D. x 20 mm, Thermo Fisher Scientific) and an Acclaim Pepmap100 C18 nano column (75 µm i.d. x 250 mm, Thermo Fisher Scientific) was operated at a flow rate of 300 nl/min and a 60 min linear gradient of 4 % to 40 % acetonitrile in 0.1 % formic acid. The LC was online-coupled to a maXis plus UHR-QTOF System (Bruker Daltonics) via a CaptiveSpray nanoflow electrospray source. Acquisition of MS/MS spectra after CID fragmentation was performed in data-dependent mode at a resolution of 60 000. The precursor scan rate was 2 Hz processing a mass range between m/z 175 and m/z 2000. A dynamic method with a fixed cycle time of 3 s was applied via the Compass 1.7 acquisition and processing software (Bruker Daltonics). Prior to database searching with Protein Scape 3.1.3 (Bruker Daltonics) connected to Mascot 2.5.1 (Matrix Science), raw data were processed in Data Analysis 4.2 (Bruker Daltonics). Swiss-Prot Homo sapiens database (release-2020_01, 220420 entries) was used for database search with the following parameters: enzyme specificity trypsin with one missed cleavage allowed, precursor tolerance 0.02 Da, MS/MS tolerance 0.04 Da, Mascot peptide ion-score cut-off 25. Deamidation of asparagine and glutamine, oxidation of methionine, carbamidomethylation or propionamide modification of cysteine were set as variable modifications.

6.14.3 *In situ* protein docking

In situ protein docking simulations were performed by Torben Fürtges and Till Rudack. To investigate the protein-protein interactions between hPol I and Top2a, the HADDOCK[213], AutoDock Vina[214], ZDOCK webserver[215], and PRISM webserver[216,217] were used. We focused on the RPA1 subunit of hPol I (PDB 7OBB, chain A) and considered both states of Top2a (state I: PDB 6ZY7, state II: PDB 6ZY8).

HADDOCK

For docking with the software package HADDOCK[213], first all amino acids of the solvent-accessible surface area were identified using FreeSASA[272]. The segment containing residues 1 060 - 1 155 of the RPA1 subunit (chain A) of the polymerase was defined as the active docking part. The surface of Top2a was defined as passive docking partner and thus completely sampled by the RPA1 subunit. Finally, the complete polymerase complex was aligned to the docked RPA1 subunit and only those docking results were considered, which do not exhibit any overlap with Top2a.

AutoDock Vina

The software package AutoDock Vina[214] was used for rigid docking of the key fragment (residues 1 060 - 1 155) of the RPA1 subunit to Top2a in both states. Finally, the complete hPol I complex was aligned to the docked RPA1 subunit, and only those docking results were considered which do not exhibit any overlap with Top2a.

Webserver docking

We used the webservers ZDOCK[215] and PRISM[216,217] using the default settings. Because of limitations of the webservers, we did not dock the complete RPA1 subunit, but (1) the key fragment (residues 1 060 - 1 155) of the RPA1 subunit (PDB 7OBB, chain A), (2) a complex of this fragment with subunit RPABC1 (PDB 7OBB, chain E), and (3) the HMG box 5 of humane UBF (PDB 2HDZ) to both states of Top2a. Finally, the complete hPol I was aligned to the docked RPA1 subunit, and only those docking results were considered, which do not exhibit any overlap with Top2a.

Protein-protein interaction analysis

The key inter-protein atomic interaction patterns were identified and analyzed using the MAXIMOBY (CHEOPS) contact matrix algorithm and the VMD plugin PyContact[218].

6.14.4 Re-analysis of previously published ChIP data sets

Re-analysis of published ChIP data was performed by Jean-Clement Mars and Tom Moss. Raw data were handled, mapping coordinates exacted, and the data displayed as previously published[221]. The used data were as follows: Top2A GSE99197_SRR5585950_TOP2A-MEF[219]. ArrayExpress E-MTAB-5839 data sets were as follows: ChIP-seq_UBF_MEFs_UBFfl_Rep1; ChIP-seq_RPI_MEFs_UBFfl_Rep1; ChIP-seq_Rrn3_MEFs_UBFfl_Rep1; ChIP-seq_TBP_MEFs_UBFfl_Rep1; and ChIP-seq_TAF68_MEFs_UBFfl_Rep1[19]. Taf1c is not included in the figure because it is identical to the Taf1b mapping, but data are also available in E-MTAB-5839 as ChIP-seq_TAF95_MEFs_UBFfl_Rep1.

6.14.5 Top2a co-immunoprecipitation

Top2a co-immunoprecipitation was performed by Konstantin Panov. To investigate Top2a interaction partners, co-immunoprecipitation was performed from U2OS Nuclear Extract (15 mg/ml total protein). Top2a was immunoprecipitated using an anti-Top2a antibody immobilized on Dynabeads Protein A magnetic beads (Thermo Fisher Scientific) according to the manufacturer's instruction. Antibodies were cross-linked to beads using DPM as recommended by the manufacturer. Beads were blocked with BSA in PBS overnight. 100 µl NE was diluted by co-IP dilution buffer to a final KCl concentration of 150 mM and treated by 500 U of benzonase for 30 min at 4°C. 25 µl of the beads were added, and the suspension was incubated on a rotating wheel for 1 h at 4°C. Beads were washed three times with 100 µl co-IP wash buffer and proteins were eluted by incubation in 1× LDS sample buffer at 65°C for 10 min. Immunoprecipitated proteins were analyzed by Western blot using anti-UBF, anti-RPA49, and anti-Top2a antibodies.

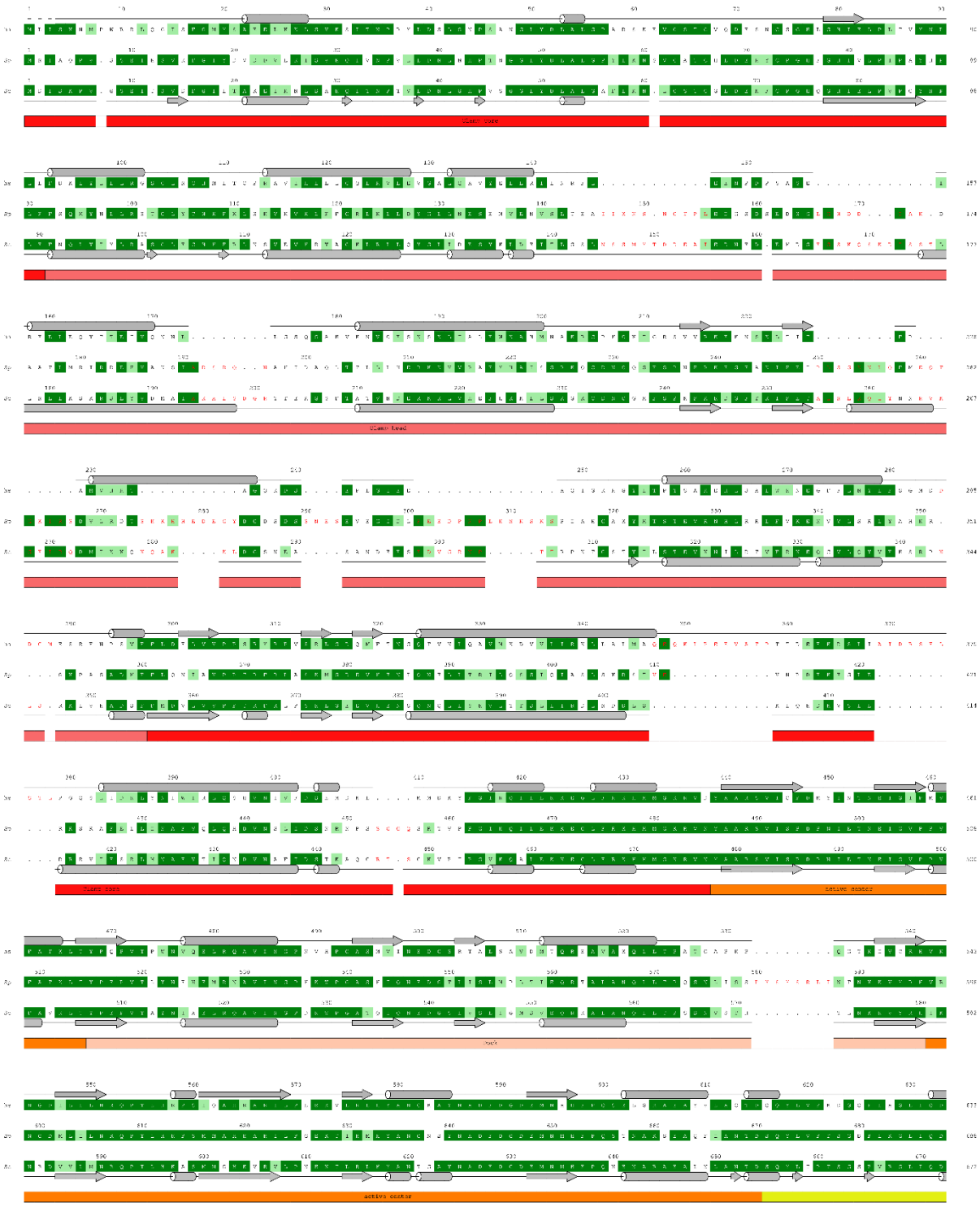
6.14.6 UBF-Top2a pulldown

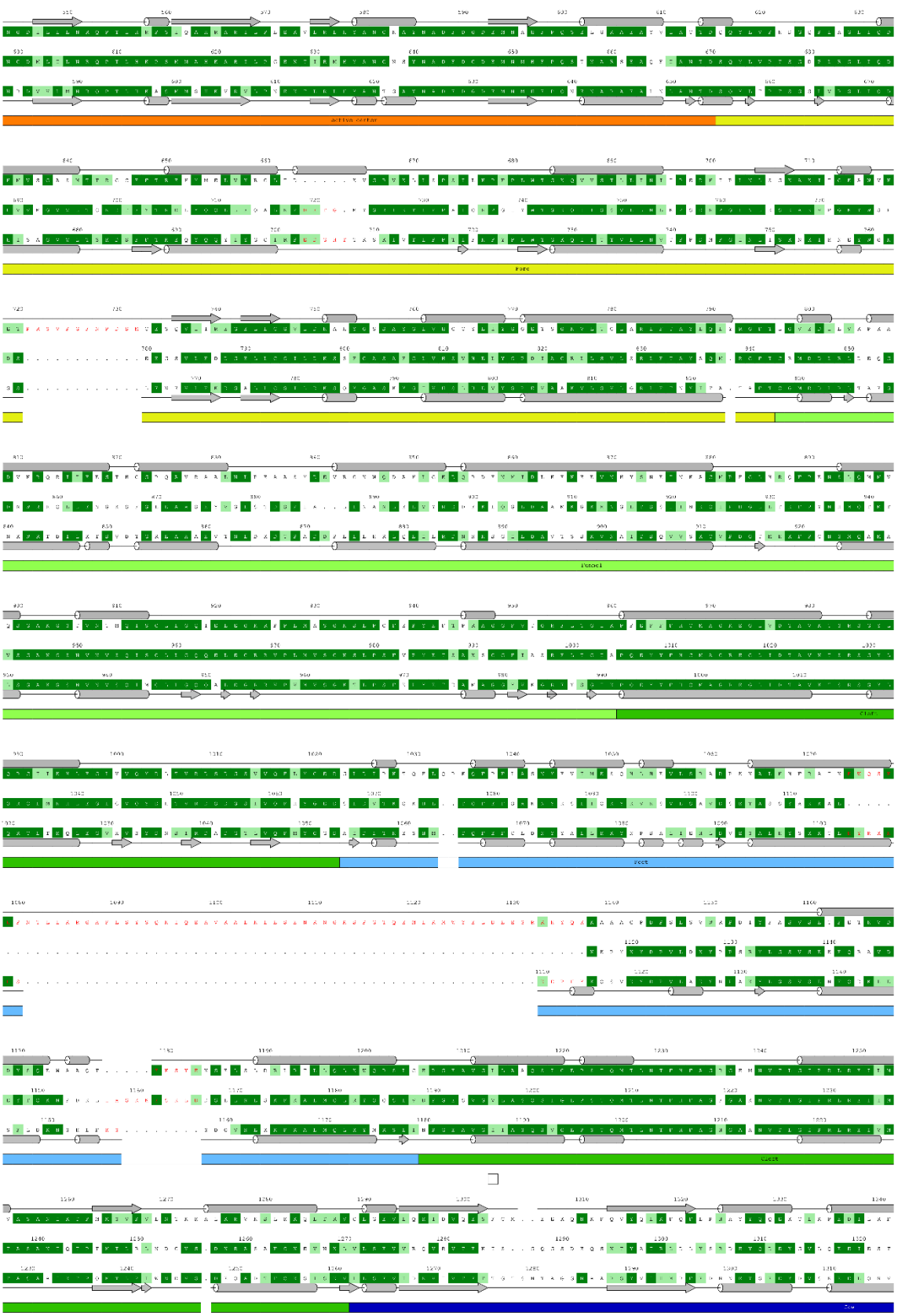
UBF-Top2a pulldowns were performed by Konstantin Panov. To investigate protein-protein interaction, a pulldown assay using purified recombinant Flag-tagged UBF (fUBF) and purified Top2a was performed. fUBF was expressed in insect cells and purified as described earlier[120]. Top2a was obtained from Inspiralis (HT210). Proteins were incubated together in UBF-Top2a pulldown buffer supplemented with 50,100, or 200 mM KCl (as marked in the figure) for 20 min at 4°C. To each sample, 20 µl anti-FLAG M2 Magnetic Beads were added and the suspension was incubated on a rotating wheel for 30 min at 4°C. Beads were washed three times with UBF-Top2a pulldown buffer supplemented with 50, 100, or 200 mM KCl, respectively, and proteins were eluted by incubation in 1× LDS sample buffer at 65°C for 10 min. Proteins were analyzed by Western Blot using anti-UBF and anti-Top2a antibodies.

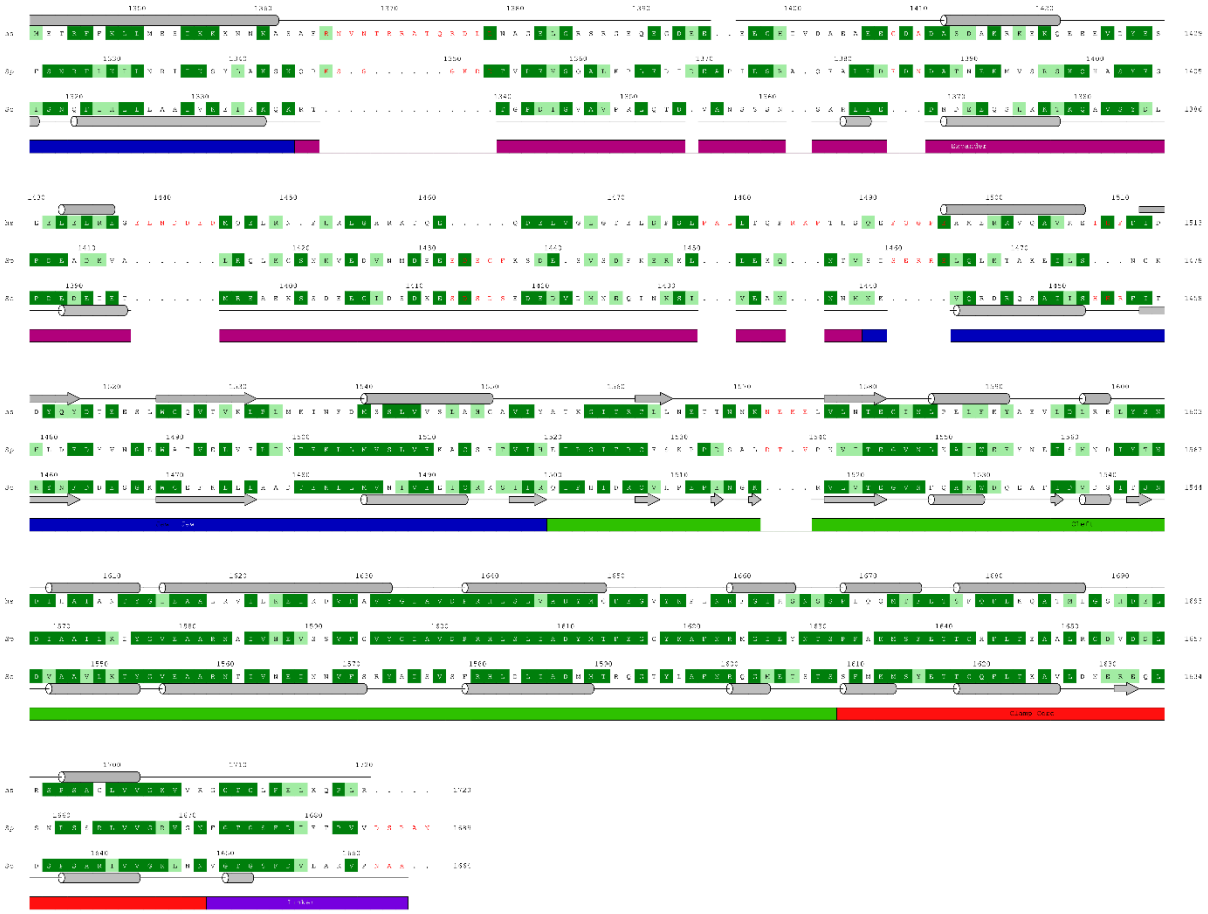
7 Appendix

7.1 Structure-based sequence alignments of Pol I subunits

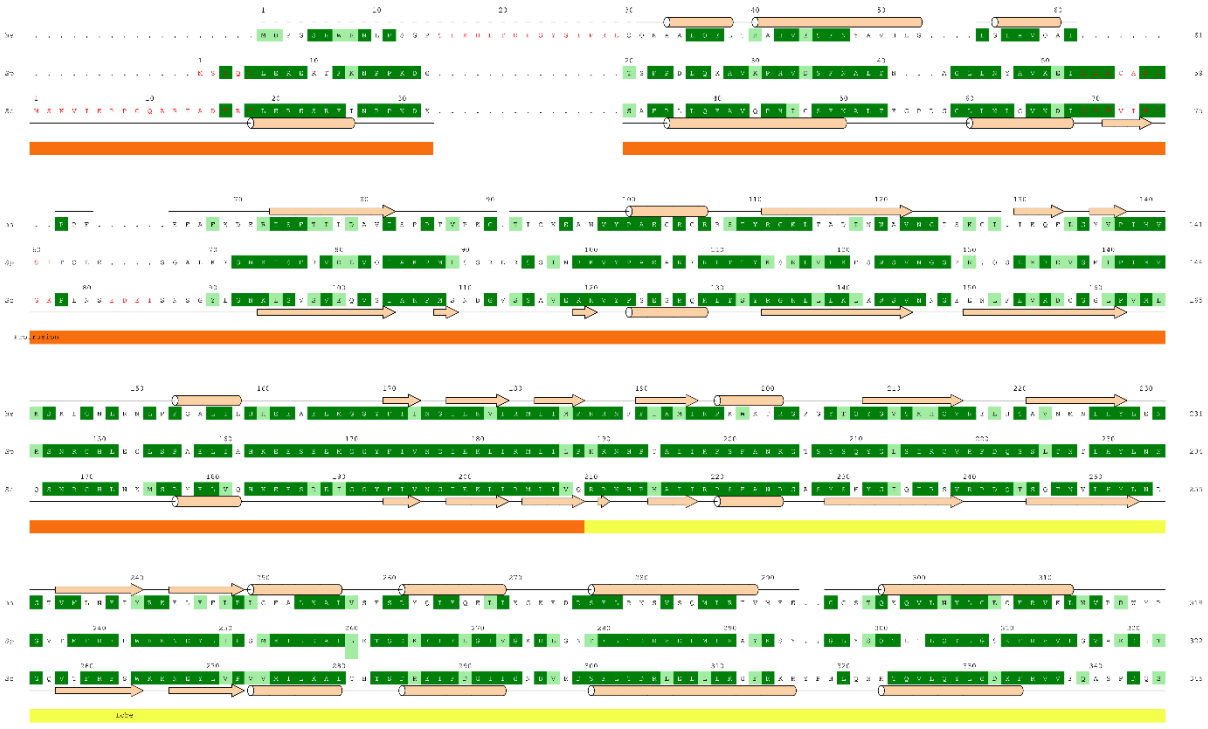
Subunit RPA1:

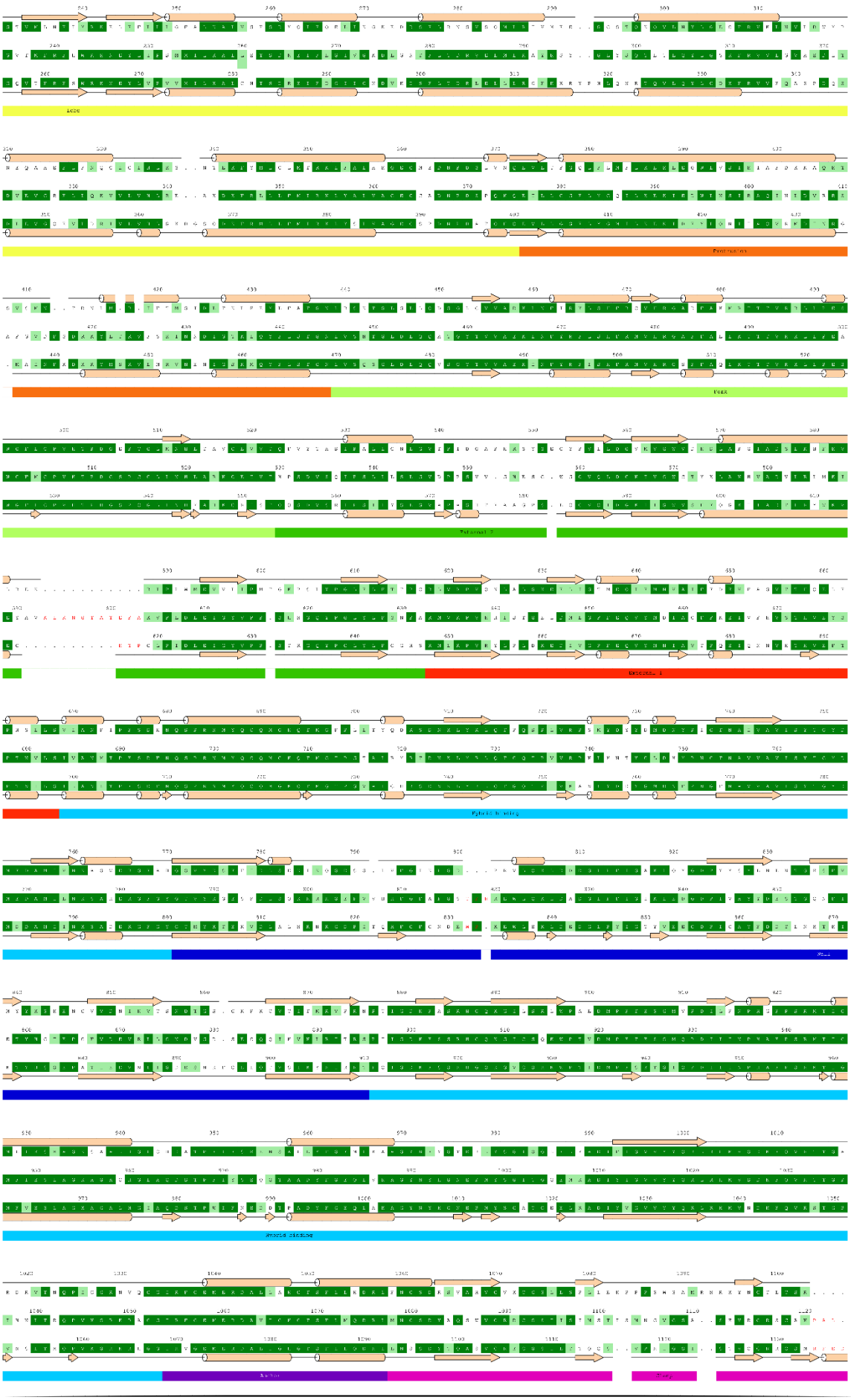


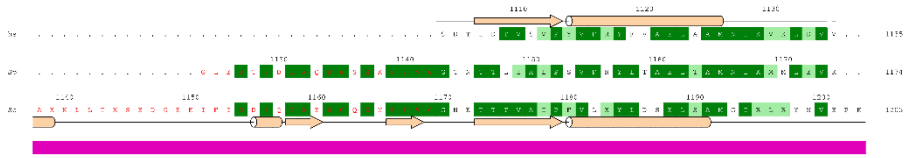




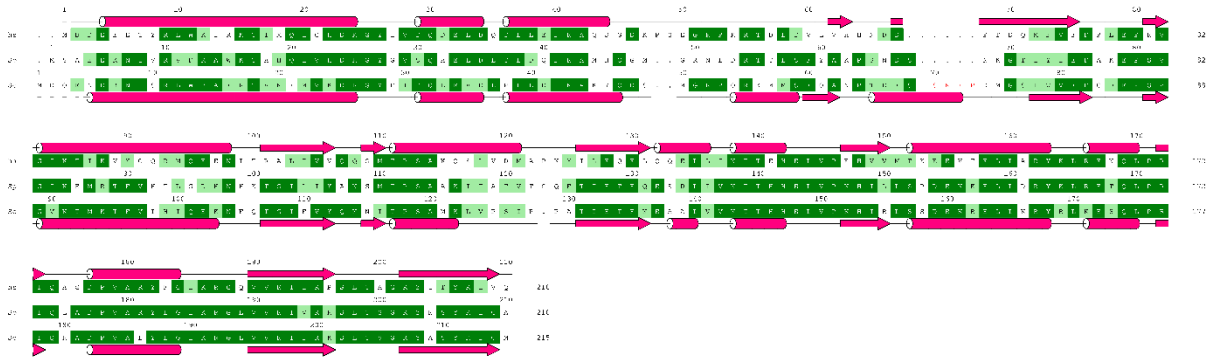
Subunit RPA2:



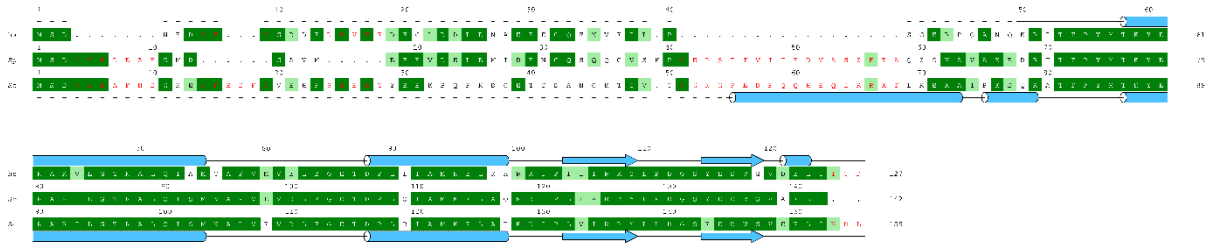




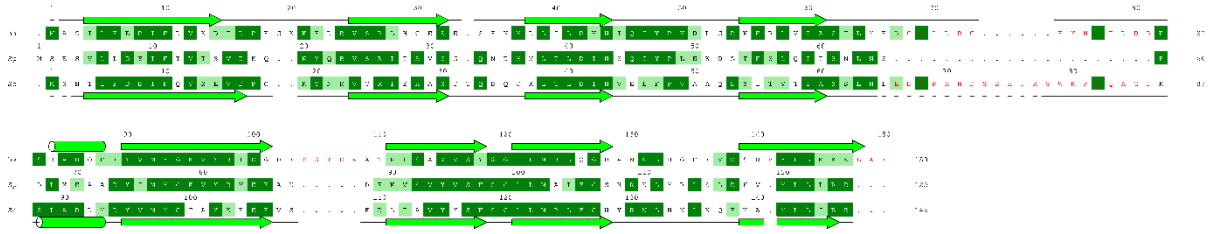
Subunit RPABC1:



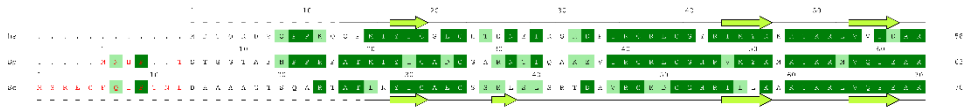
Subunit RPABC2:



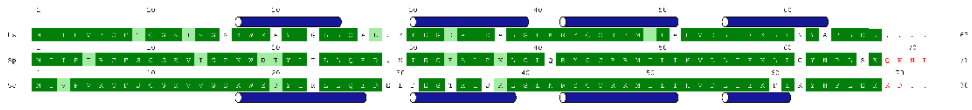
Subunit RPABC3:



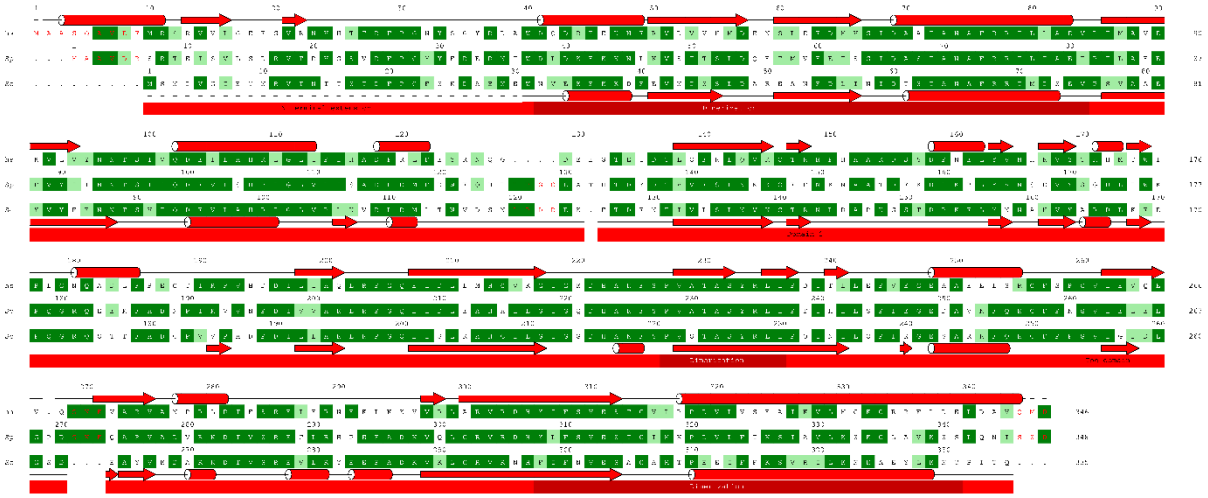
Subunit RPABC4:



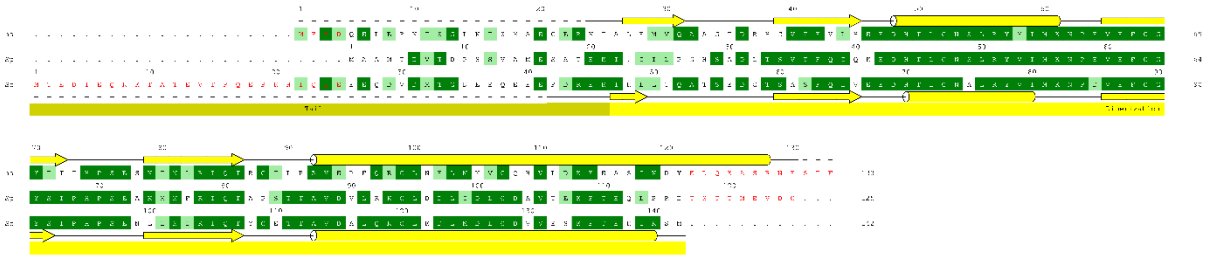
Subunit RPABC5:



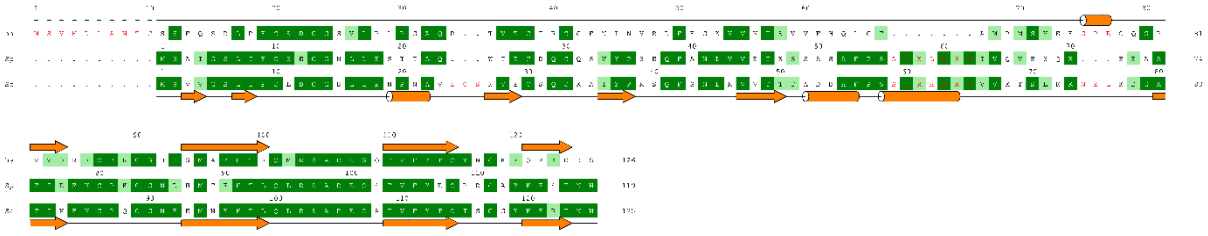
Subunit RPAC1:



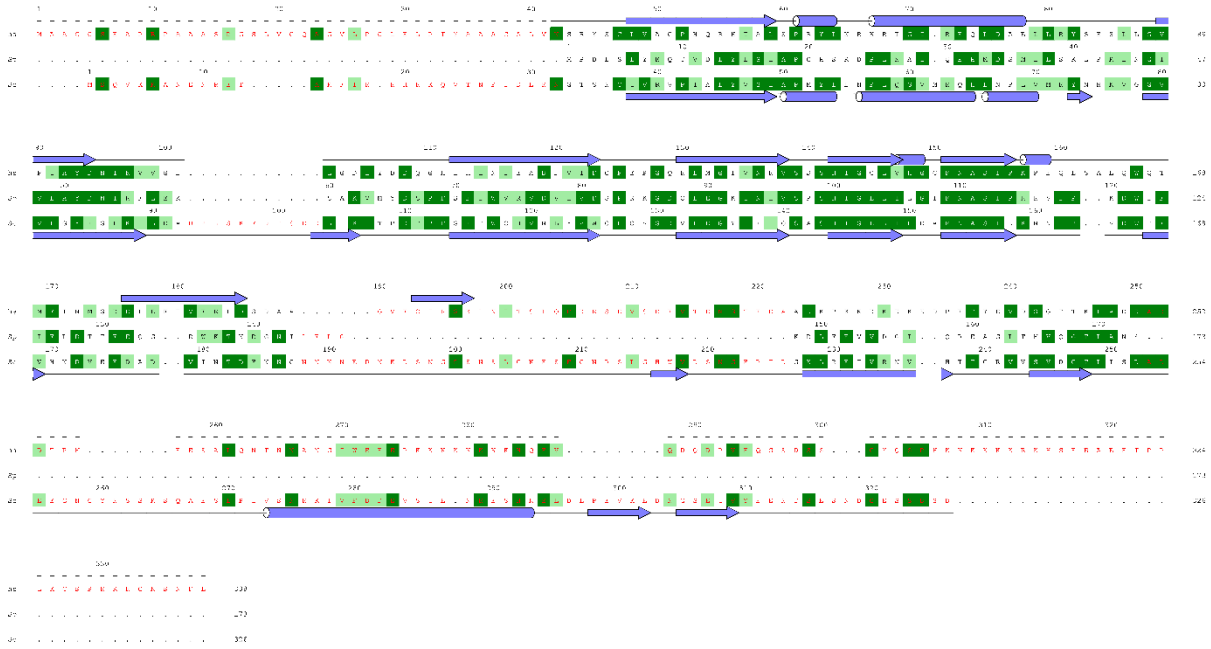
Subunit RPAC2:



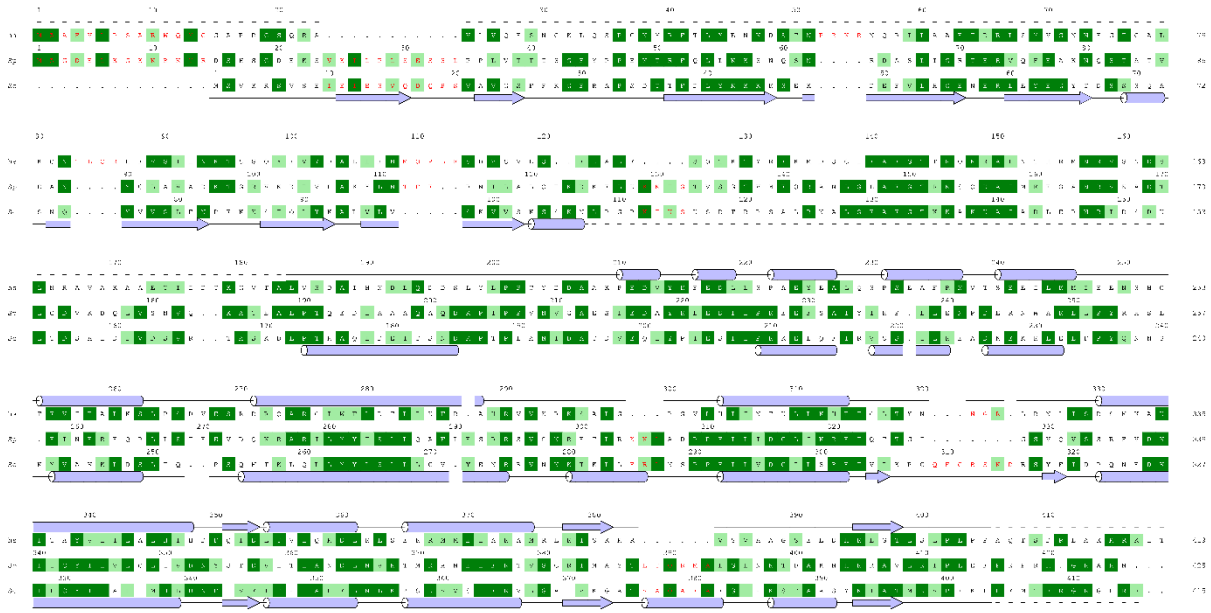
RPA12:



RPA43:



RPA49:



Subunit RPA34:

7.2 Structure-function investigations of human Pol III

7.2.1 Introduction

Pol III is the largest of the three DNA-dependent RNA polymerases (Pols) in the eukaryotic system and is responsible for the transcription of many non-coding RNAs as tRNAs, 5S rRNA, U6 RNA, 7SK RNA and others[2,9,273]. As these RNAs are essential for precise gene transcription and protein translation, it's not surprising that the Pol III transcription is directly linked to cell proliferation and is a determinant of eukaryotic lifespan[274]. As a consequence, Pol III is cell cycle- and cell type-dependent regulated and deregulation is found in various cancers[275–277]. Many known mutations within Pol III lead to different diseases, which often include a neurodegenerative phenotype[198–200,278–282]. Moreover, some mutations within Pol III cause a hypersensitivity to viral infections[283–285].

RNA polymerase III

The Pol III enzyme itself is highly conserved and consists of 17 subunits[7]. Ten subunits are unique for Pol III (RPC1-10), the five common subunits (RPABC1-5) are built into all three polymerases and subunits RPAC1 and RPAC2 are shared with Pol I (see Pol III overview Figure 26A)[7,137,286]. RPC1 and RPC2, homologous to RPA1 and RPA2 in Pol I, are the two largest subunits, build up most of the 10-subunit core enzyme, and are responsible for most of the catalytic activity. The protruding stalk subcomplex of Pol III is composed of two subunits (RPC8 and RPC9) and is important for transcription initiation. It was shown that the stalk binds tRNAs with high affinity, and might be able to bind newly synthesized RNAs just after leaving through the RNA exit channel[287,288]. Subunit RPC10 includes the intrinsic cleavage activity with its C-ribbon ensuring transcription fidelity and proof-reading activity, binds to the lobe with its N-ribbon and is essentially homologous to subunit RPA12 of Pol I[52,289–291]. The C-ribbon together with the linker region was shown to be important for Pol III reinitiation-recycling[292], and the N-ribbon is indispensable for proper transcription termination[293]. The lobe-binding hetero-dimer RPC4/5, homologous to RPA49/34 of Pol I and TFIIF of Pol II[294], can be seen as a built-in transcription factor. It is necessary for proper transcription initiation as well as termination[295–297]. The C-terminal extension of RPC5 is not conserved among organisms and additionally enlarged in the human enzyme compared to yeast, similar to the situation of RPA34 in humans. The heterotrimer RPC3/6/7 of Pol III is another built-in transcription factor complex only present in Pol III and is distantly related to the Pol II transcription factors TFIIEα and TFIIEβ[7].

Pol III promoters

In contrast to Pol I, Pol III needs to recognize three different types of promoters, which is still a lot less diverse than Pol II promoters (Figure 26B)[298,299]. Type I promoters, which are present at 5S rRNA genes, are characterized by an internal control region (ICR), which gets bound by initiation factors TFIIIA and TFIIIC, which again recruit TFIIIB[22,298,300]. In type II promoters, present in tRNA genes, TFIIIC directly binds to the conserved A and B boxes and recruits TFIIIB[22,298]. ICR, A box, and B box are located down-stream of the transcription start site (TSS), whereas the TBP-containing factor TFIIIB binds upstream of the TSS[301]. Type III promoters are only present in vertebrates and occur in front of many different genes, such as U6 or 7SK RNA genes[22]. Type III promoters contain a TATA-box motif, which is recognized

by TFIIIB and a proximal sequence element (PSE) bound by the snRNA activating protein complex (SNAPc)[22,302,303]. Here, the TATA-box is located upstream of the TSS, and the PSE even more upstream, again allowing a direct protein-protein interaction between the two initiation factors TFIIIB and SNAPc[301].

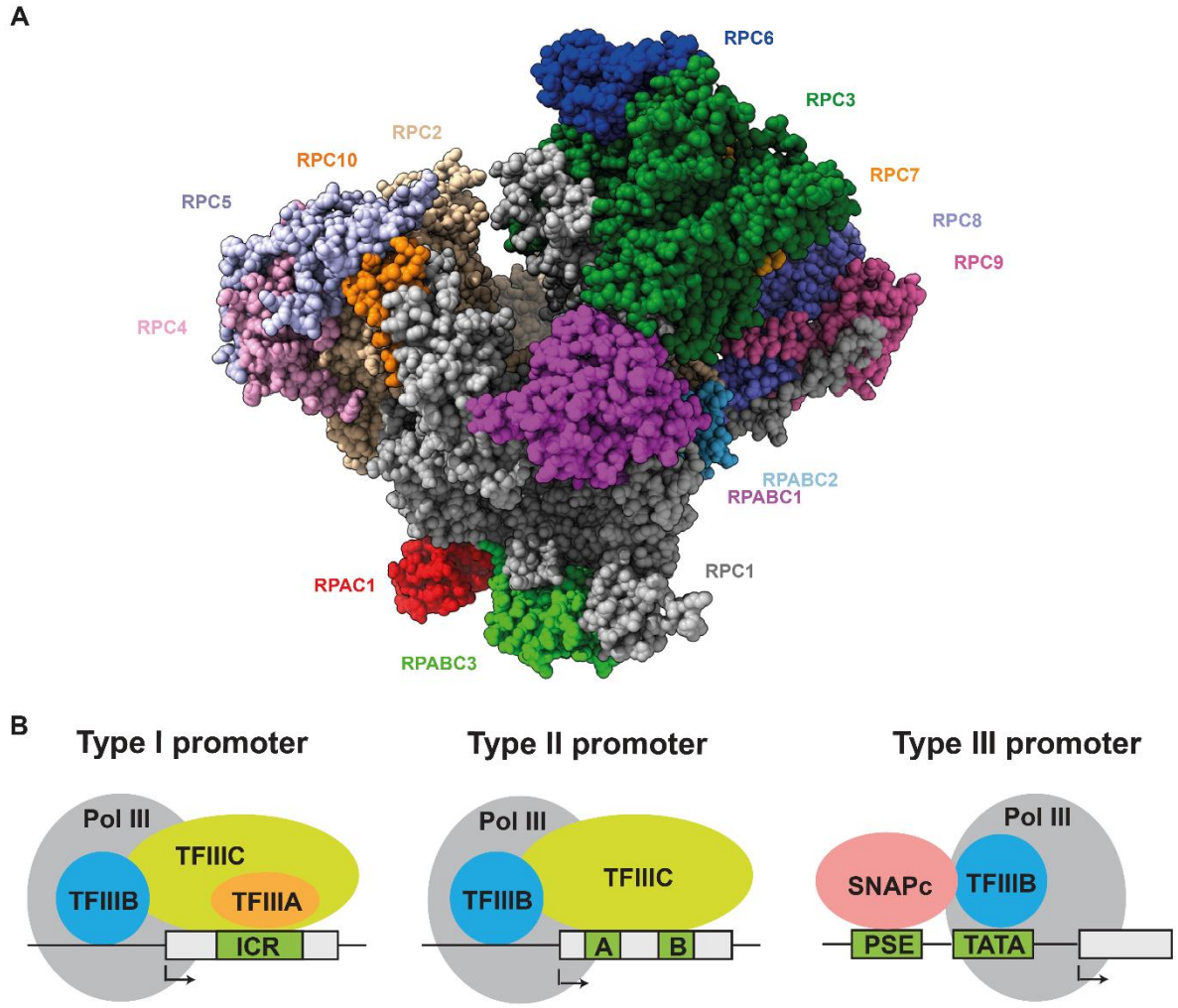


Figure 26 Pol III architecture and promoters

A Structure of monomeric *S. cerevisiae* Pol III (PDB 6EU2)[304] with colored subunits: RPC1: grey; RPC2: wheat; RPAC1: red; RPAC19: yellow; RPABC1: magenta; RPABC2: hafnium; RPABC3: green; RPABC4: lemon; RPABC5: density; RPC8: slate; RPC9: hot pink; RPC10: orange; RPC5: light blue; RPC4: pink; RPC3: dark green; RPC6: dark blue; RPC7: light orange. **B** The composition of initiation complexes formed on type I, II, and III promoters are schematically illustrated. Black arrows mark the transcription start site.

Pol III initiation factors and regulatory protein MAF1

Single-subunit TFIIIA contains 9 zinc finger motifs, binds Type I promoters, and has a high affinity towards the ICR of the promoter[305]. Similar to Pol III, it can bind its gene product, 5S rRNA and hence plays a role in 5S rRNA gene expression and cytoplasmic export[22,306–308].

The trimeric TFIIIB complex is composed of TBP, Bdp1 (B double prime 1) and Brf1 or Brf2 (B-related factor 1 or 2) and binds to all three promoter types. Besides the TBP-DNA interaction, Bdp1 binds to

the major and minor groove of the promoter DNA and enhances the TFIIIB-DNA complex stability[309]. In each TFIIIB complex either Brf1 or Brf2 is included. In case Brf2 is present, TFIIIB binds to type III promoters [310]. Additionally, Brf2 functions as a redox-sensor and modulates Pol III transcription adjusted to the cellular oxidative stress responses[310].

Six-subunit TFIIIC is composed of the two subcomplexes τA (TFIIIC102, TFIIIC63, TFIIIC35) and τB (TFIIIC220, TFIIIC110, TFIIIC90), which are flexibly connected and bind to the A and B boxes of type II promoters, respectively[302,311–313]. Via τA a direct protein-protein interaction with TFIIIB is formed and helps regulation of promoter recognition and transcription initiation[314].

SNAPc, important for type III promoter transcription initiation, is composed of five subunits (SNAP43, SNAP45, SNAP50, SNAP190, SNAP19) and only exists in vertebrates[301,302]. This initiation factor interacts DNA-independent with the TFIIIB initiation factor, but most of the action of SNAPc stays elusive up to now[302,309].

Besides these initiation factors, which are mostly composed of multiple subunits (except for TFIIIA), MAF1 is a major regulator of Pol III transcription[315,316]. This single-subunit protein negatively regulates Pol III transcription of 5S rRNA and tRNAs, nevertheless some genes (including specific tRNAs) are less sensitive to MAF1-regulation[317,318]. MAF1 negatively regulates Pol III by direct binding, thereby preventing Pol III - TFIIIB interaction and consequently repressing initiation[319]. MAF1 itself is regulated by phosphorylation, including PKA, Sch9, casein kinase II, and TORC1 which phosphorylate and thereby inactivate MAF1, leading to increased Pol III activity[320–322]. Contrary, PP2A (phosphatase type 2A) dephosphorylates MAF1, inducing a repression of Pol III transcription[323].

Scope of this project

Similar to the hPol I project, the aim was to allow transfer of the knowledge from yeast Pol III to the human enzyme by investigating which mechanisms are conserved or divergent. Comparable to the situation in Pol I research, yeast Pol III has been studied in detail using purified and recombinantly expressed components allowing a detailed description of subunit (sub-)domain and transcription factor functionalities throughout the transcription cycle. In higher eukaryotes, the demand of accurate and detailed regulation is increased as a result of additional layers of regulation in a multi-cellular organism. Different mutations in Pol III subunits can lead to various neurodegenerative diseases or impair immune response and misregulation of hPol III transcription has been associated with many types of cancer.

In this study, I aimed to investigate the conserved and divergent mechanisms of Pol III transcription between yeast and human by a detailed structure-function analysis. For this purpose, a well-defined *in vitro* system was established consisting of highly purified components. A human cell line was modified in order to allow purification of the multi-subunit Pol III complex. Pol I and Pol III shared subunit RPAC1 of this cell line is carrying a protease-cleavable sfGFP-tag enabling the purification of endogenous hPol I and hPol III to high purity and quality in a close-to-native state. Purified hPol III is further used for a detailed structure-function analysis including the cryo-EM reconstruction of its monomeric state. The structure now allows mapping and evaluation of the effects of disease-causing mutations and will support anti-cancer drug development.

7.2.2 Results

HeLa-POLR1C cell line

Comparable to the HeLa-POLR1A cell line, the cell line HeLa-POLR1C was first established using the same strategy as described in the method section 6.5.7 for HeLa-POLR1A. This allowed for the purification of human Pol I and Pol III at the same time as subunit RPAC1, shared between the two enzymes, was tagged. The parental cell line HeLa P2[180] was modified with the help of a CRISPR/Cas9 double-nickase strategy and a homology template allowing homology-directed repair (HDR) and hence the introduction of a sfGFP-tag (Figure 27A). The double-nickase strategy with two gRNAs, resulting in a staggered double-strand break in close proximity to the stop codon of RPAC1, was used to minimize potential off-target effects. The introduced sequence just in front of the endogenous stop codon is composed of a GS-linker including a 3C protease cleavage site and the ORF of sfGFP. Several days after transfection of the cells, GFP-positive cells were selected and individualized by FACS. Afterwards, the monoclonal clones were tested for homozygous insertion of the DNA-sequence within the genome by Western Blot (Figure 27B). A positive clone with homozygous insertion was selected and the newly established cell line was named HeLa-POLR1C.

Confocal microscopy was used to check complex integrity and to visualize the cellular GFP, and thereby examining Pol I and Pol III localization (Figure 27C). Bright dots of GFP-signal were found in the nucleolus going along with the tagged Pol I enzyme. Additionally, more distributed background within the nucleus was present which corresponds with the distributed localization of Pol III. Furthermore, some brighter dots were visible in the cytoplasm, which might represent cytosolic Pol III, which is involved in viral defense mechanisms within the cytoplasm[324]. As the cells were characterized by a normal cell growth and correct localization of the enzymes, the integrity of both complexes appeared to be unchanged and the C-terminal tagged subunit RPAC1 was normally integrated within the protein complexes.

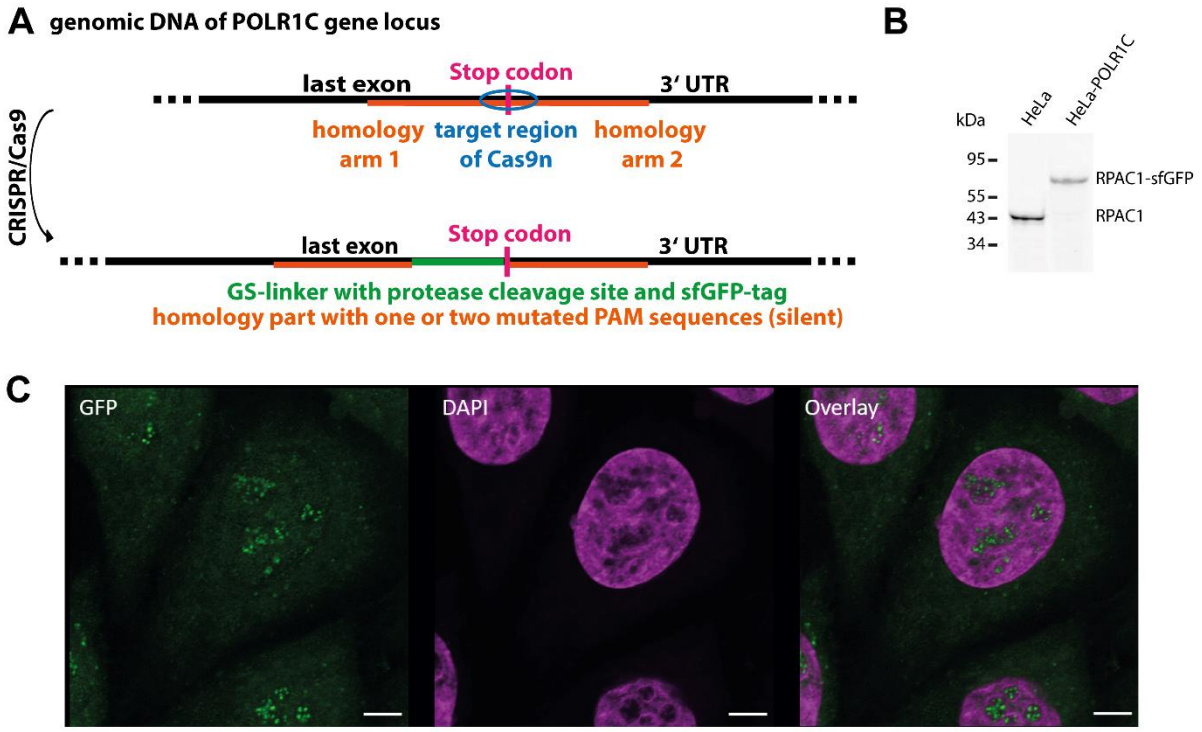


Figure 27 Human cell line HeLa-POLR1C

A The schematic overview of the used CRISPR/Cas9 strategy for the introduction of a protease-cleavable sfGFP tag (green) just before the endogenous stop codon (pink) is shown. The target region of Cas9n cleavage was in the close surrounding of the stop codon (blue) and HDR was enabled by transfected template DNA containing the two homology arms (orange) and the tag sequence (green). **B** Western blot against subunit RPAC1 showed a shift to larger molecular weight in lysates of the HeLa-POLR1C cell line compared to the parental HeLa cell line, confirming exclusive expression of the modified protein. **C** Confocal imaging of modified HeLa-POLR1C cell line expressing homozygous sfGFP-tagged RPAC1. Endogenous GFP signal, representing Pol I and III (green), DAPI staining (magenta) and overlay of both channels are shown. Scale bars: 5 µm.

hPol III purification and in vitro activity

Having the HeLa-POLR1C cell line at hand, affinity purification of hPol I and hPol III was possible. After large scale cell cultivation and harvest, the purification of Pol I and Pol III was performed analogous to the described hPol I purification from the HeLa-POLR1A cell line (Figure 28A, see method section 6.6). After cell lysis a GFP-pulldown was conducted to trap hPol I and hPol III on the beads. Elution was executed by protease cleavage with 3C protease in order to separate the bead-bound sfGFP-tag from the polymerases. The eluate consisting of 3C protease, hPol I, and hPol III was further separated by anion-exchange chromatography. 3C protease was found in the flow-through, whereas hPol I elutes at lower salt concentrations than hPol III. Although, hPol I elution fractions were always slightly contaminated with hPol III, hPol III fractions showed a high degree of purity. SDS-PAGE and mass spectrometry analysis revealed that all 17 Pol III subunits were present (Figure 28B).

Having *in vitro* purified hPol III allowed for structure-function analysis. First, *in vitro* catalytic activity was tested with the same assays as conducted for hPol I (see method section 6.12.2). A minimal transcription scaffold was used, composed of a short double-stranded DNA with a single-stranded 3' overhang of the template strand to which an RNA primer was annealed. Titrating freshly purified hPol III into the reaction with this transcription scaffold showed cleavage (without NTPs) and elongation (with

NTPs) activity (Figure 28C). This verified *in vitro* activity of purified hPol III which was further used for structural investigations by cryo-EM.

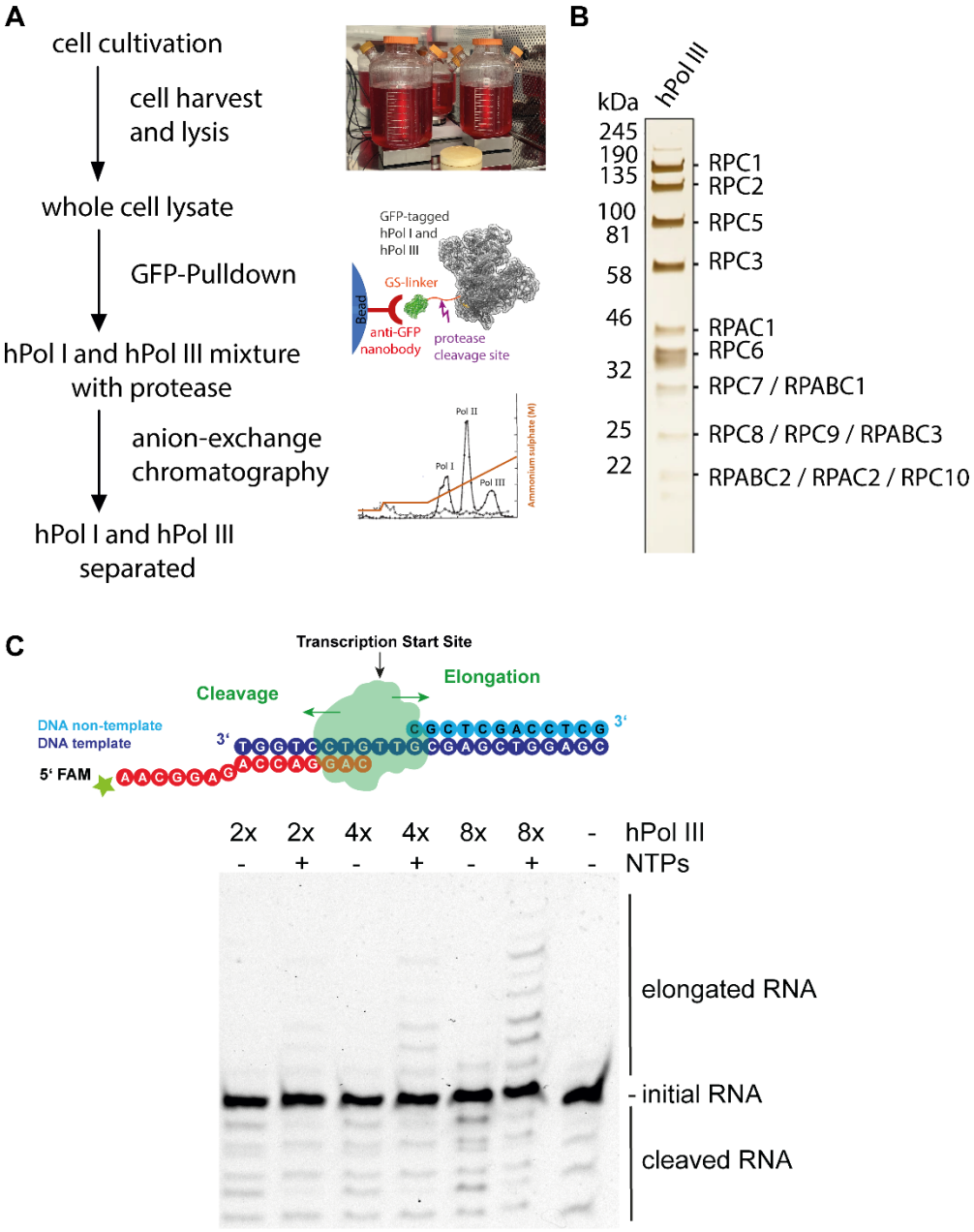


Figure 28 Purification and transcription of hPol III

A A schematic representation of the hPol I and hPol III purification protocol is shown. (PDB 1GFL for GFP, 6EU2 for Pol III; anion-exchange chromatography from Roeder and Rutter, 1969[2]) **B** Silver-stained SDS-PAGE of purified human Pol III shows bands for all subunits. **C** On top a schematic representation of the assay and scaffold sequence used to determine hPol III activity *in vitro* is shown. Underneath, RNAs detected by their fluorescent label, separated by length and an Urea-PAGE sequencing gel showed cleavage (without NTPs) and elongation (with NTPs) activity of hPol III with increasing enzyme amounts (2x, 4x and 8x molar access compared to the DNA/RNA scaffold). As a control, scaffold-only was treated identically without any Pol for reference.

hPol III structure

Freshly purified hPol III was first diluted to 110 mM ammonium sulphate and concentrated to 0.06 mg/ml for structural analysis. The sample was applied to carbon-coated grids and vitrified. As hPol III possessed a biased orientational distribution on the carbon support film, two datasets with different tilt angles (0° and 30° tilted) were collected on a Titan Krios TEM microscope equipped with a Falcon III camera to overcome this problem of biased orientational distribution. The merged dataset was processed with MotionCor2[325], CTFFIND4[326], cisTEM[327], and the RELION 3.1 suite[190]. 2D classification (172 678 particles for merged datasets) and hierarchical 3D classifications (25 369 particles in final reconstructions) were performed to select a stable set of particles (Figure 29, Table 6). The resulting cryo-EM density of the non-crosslinked apo polymerase showed an overall resolution of 4.0 Å. The majority of hPol III molecules showed the intact 17-subunit enzyme including the heterotrimer RPC3/6/7, but there was also a sizeable fraction missing the heterotrimer (Figure 30). This could be caused by a partial loss during purification, which would be in line with earlier reports[328] or could have happened during cryo-EM specimen preparation. Whereas the core polymerase possessed more details, stalk (RPC8/9), heterotrimer (RPC3/6/7), and parts of the clamp showed lower local resolution and more flexibilities. Taken together, these observations suggest, that the coiled-coil region of the clamp and the heterotrimer, which are contacting each other, build a distinct structural and functional unit together. Additionally, the C-terminal extension of RPC5, which was shown to be metazoan-specific and not present in yeast, stayed among the flexible parts and cannot be resolved.

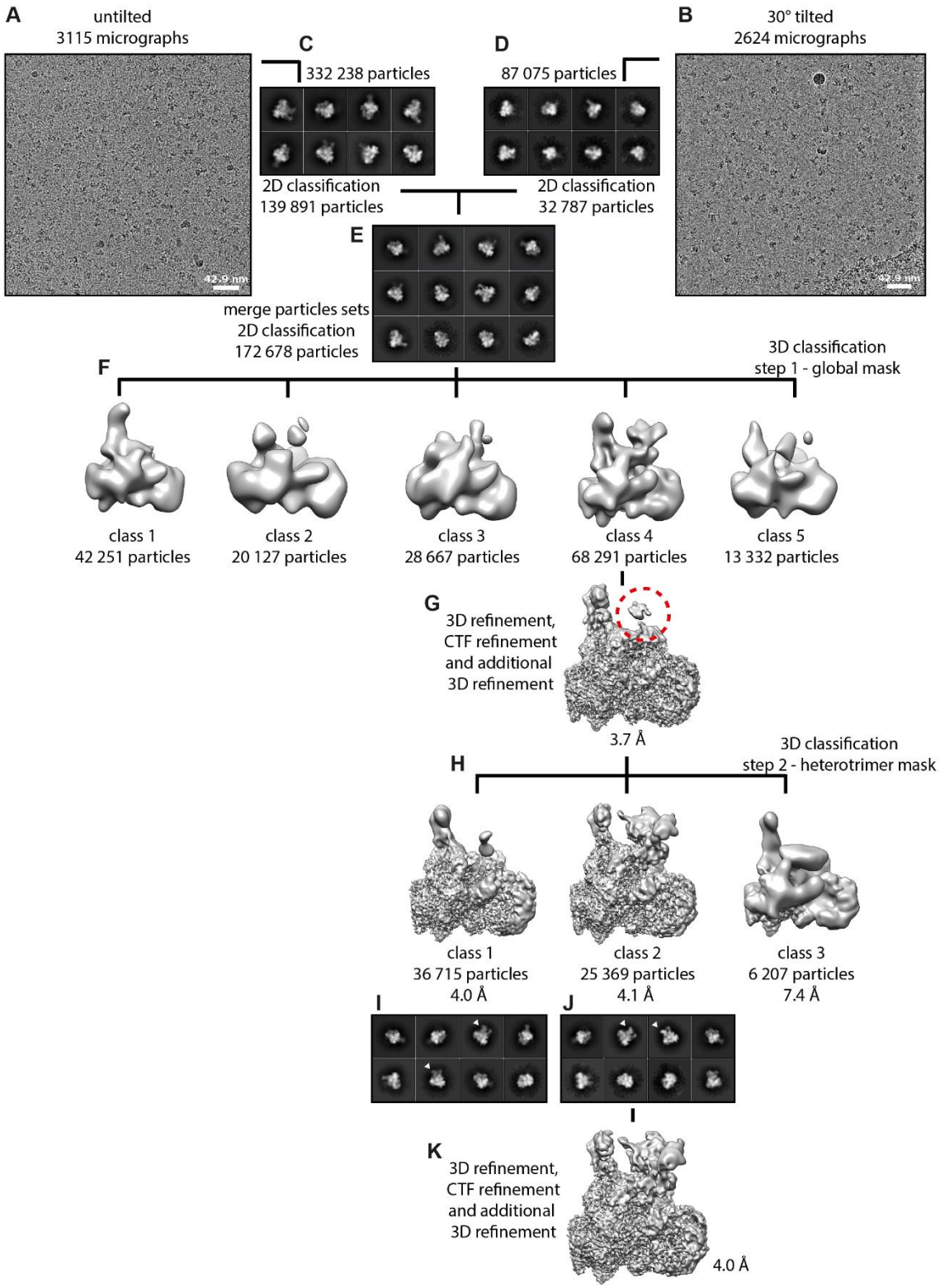


Figure 29 Processing tree of hPol III cryo-EM reconstruction

A, B Representative micrograph for untilted (A) and 30° tilted (B) datasets. **C, D, E** Representative 2D class averages for untilted (C), 30° tilted (D) and merged (E) datasets. **F** Consensus 3D classification of the resulting particle set using a Cryosparc *ab inito* model as reference. **G** The class corresponding to full polymerase was refined. **H** Masked classification around the heterotrimer (marked with red circle). **I, J** Representative 2D classes of classes from focused 3D classification, with the heterotrimer density marked. **K** The class corresponding to the full polymerase molecule was subsequently refined to generate the final model.

Table 6 Cryo-EM data collection and refinement statistics of hPol III

	Human Pol III
PDB	7AST
EMDB	11904
Data collection (0° Tilt/30° Tilt) and processing	
Voltage (kV)	300/300
Electron exposure ($e^-/Å^2$)	44.1/37.8 and 40.6
Defocus range (μm)	-1.0 - -3.0/-1.2 - -3.0
Pixel size (Å)	1.065/1.065
Initial particles images (no.)	172,678
Final particle images (no.)	25,369
Map resolution (Å)	4.0
FSC threshold	0.143
Map sharpening *B* factors ($Å^2$)	-116.839
Refinement	
Model composition	
Non-hydrogen atoms	34,636
Protein residues	4,369
Nucleotides	-
Ligands	-
R.m.s. deviations	
Bond lengths (Å)	0.016
Bond angles (°)	1.086
Validation	
MolProbity score	2.74
Clashscore	31.792
Poor rotamers (%)	0.0
Ramachandran plot	
Favored (%)	78.66
Allowed (%)	20.94
Disallowed (%)	0.39

Overall, the structure of Pol III is highly conserved between yeast and human as expected from the high sequence conservation. Nevertheless, some human-specific features could be observed, which might be important for different human-specific functions, assembly, or regulation. For instance, the human heterodimer RPC4/5 had a weaker association to the core polymerase, as one helix of RPC4, which contacts RPC2 in yeast, was deleted in humans. Furthermore, the dimerization domain of RPC5 was slightly altered, which hinted to a slightly altered dimerization module of the heterodimer in humans. Moreover, the C-terminal extension of RPC5 (RPC5EXT) was only present in humans, but stayed flexible in the reconstruction, still contributing to the fact that the heterodimer differs between yeast and human. In human Pol III, two deletions were observed in the foot domain of RPC1, which lead to a rearrangement of the whole foot domain to a more compact structure. Interestingly, alterations in the foot domain were found for all three mammalian polymerases compared to their yeast counterparts[222–224,229]. In Pol II, it was shown that this domain is responsible for many protein-protein interactions and is a point of regulation[329,330], which suggests diverse or additional regulation and possibly accessory proteins in the human system.

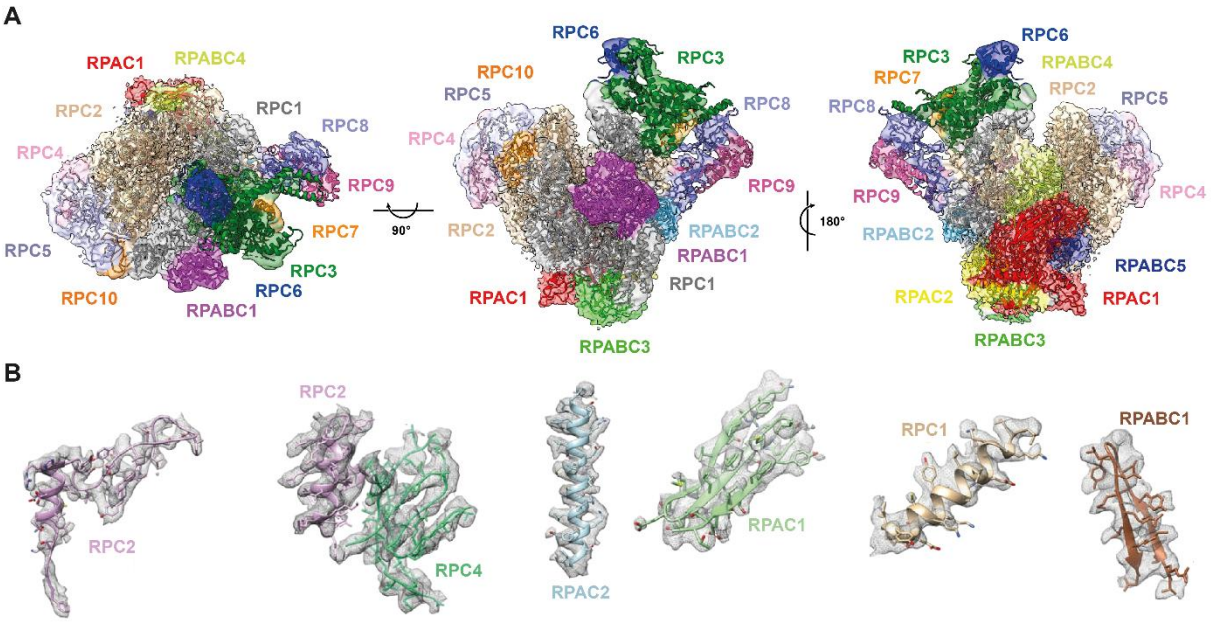


Figure 30 Structure of hPol III

A The electron density map is shown, filtered according to the local resolution with the fitted model shown in ribbon representation. Regions of the electron density map are colored according to the subunit structure: RPC1: grey; RPC2: wheat; RPAC1: red; RPAC19: yellow; RPABC1: magenta; RPABC2: hafnium; RPABC3: green; RPABC4: lemon; RPABC5: density; RPC8: slate; RPC9: hot pink; RPC10: orange; RPC5: light blue; RPC4: pink; RPC3: dark green; RPC6: dark blue; RPC7: light orange. **B** Selected regions of several subunits are shown, showing the fit with the filtered electron density (mesh).

Mapping of disease-causing mutations of hPol III

Many mutations within the hPol III subunits RPC1, RPC2, RPAC1, and RPAC2 cause various human diseases often leading to a severe phenotype (Table 7). Having now the molecular structure of hPol III allowed mapping and evaluation of these disease-causing mutations (Figure 31, Table 7). Examples for Pol III-associated diseases are Treacher-Collins syndrome (TCS)[198,199], Hypomyelinating Leukodysthrophy (HL)[200,202,280–282,331–334], and Wiedemann-Rautenstrauch syndrome (WRS)[278,279] leading to severe neurodevelopmental disorders, as well as an acute severe response to the Varicella zoster virus (VZV)[283,285,335], which is characterized by an impaired immune system in patients due to Pol III mutations.

The WRS is characterized by different mutations within the largest subunit RPC1, many of which seemed to avoid proper interaction with RPABC1. The two neurodegenerative diseases HL and TCS are both characterized by many mutations affecting hPol I and hPol III and cannot be classified as Pol I- or Pol III-associated phenotypes *per se*. Causative mutations of HL were found in subunits RPC1, RPC2, and RPAC1. Most of the mutations within the Pol III-specific subunits RPC1 and RPC2 seemed to have a destabilizing effect on intra- und inter-subunit interactions, which might cause instability of the whole enzyme. Mutations in the shared subunit RPAC1 were present in hPol I and hPol III, but some of them (T26I, T27A, P30S, N32I, N74S, I105F, H108Y, R109H/C) seemed to have a Pol III-specific effect, as they built inter-subunit interactions with Pol III-specific subunits or the regions stayed flexible in hPol I structures (compare result section 2.6.2)[222,224,280]. In line with this, mutations N32I and N74S in RPAC1 were shown to inhibit Pol III assembly and nuclear import, but do not affect Pol I[200]. TCS, another disease affecting Pol I and Pol III, is caused by mutations within RPA2, RPAC1, RPAC2, and the

Pol I regulating gene *TCOF1*. The mutations within RPAC2 seemed to destabilize directly or indirectly the assembly with RPAC1 and hence RPAC1/2 interaction with the core polymerase, whereat mutations within RPAC1 (R279Q/W) seemed to destabilize the subunit itself intrinsically, but seemed to have a more intense effect in Pol I as the corresponding region additionally mediates interaction with RPA34 (compare result section 2.6.2)[222]. Overall, both diseases, TCS and HL, appeared to be caused by defects within Pol I and Pol III and cannot be classified *per se*, although some mutations of shared subunits seemed to have an enzyme-specific effect. Thus, HL might be a Pol III-associated and TCS a Pol I-associated disease in some cases, not excluding a low extent of disability within the other polymerase and not excluding the possibility that some mutations resulting in both diseases affect both polymerases.

Causative mutations within Pol III-specific subunits (RPC1, RPC3, RPC6) were found in patients affected by an acute severe response to a Varicella zoster virus (VZV) infection[283–285,335]. All these mutations were solvent exposed within the structure and resulted in neutralizing charged residues. These residues were located in proximity to DNA-binding surfaces, which might explain the defect in cytosolic DNA-sensing of Pol III which is included in the immune response against VZV[324]. The reason, why nuclear Pol III transcription is not influenced by these mutations, might be a buffering effect of the transcription factors, which made direct DNA-binding of the polymerase unnecessary.

Overall, these findings are in line with results from mapping of disease-causing mutations on yeast structures[280–283,336], but further evaluated their effects and could promote further understanding of the disease mechanisms as well as drug development.

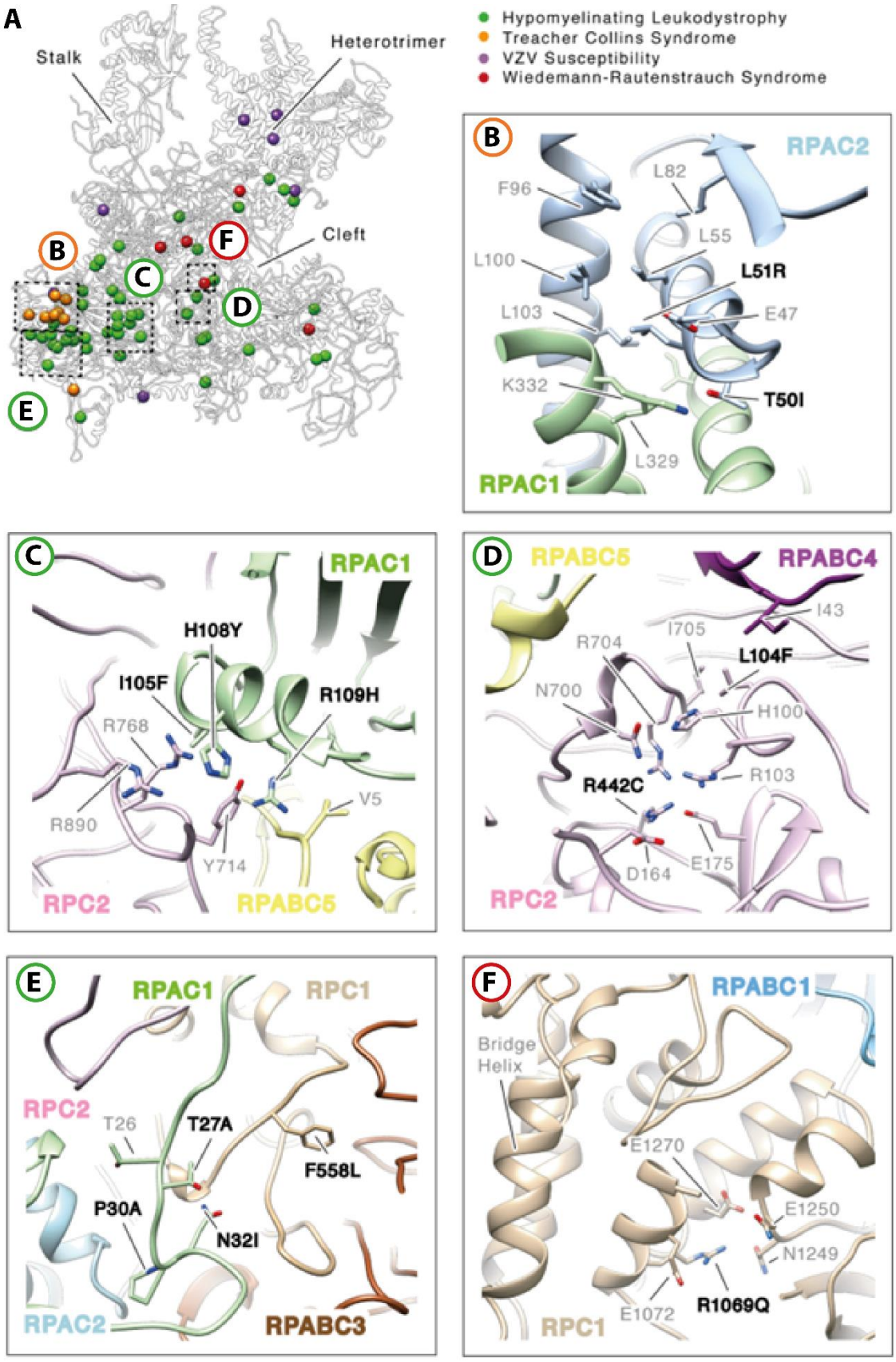


Figure 31 Mapping of disease-causing mutations with hPol III

A Structural model of hPol III (white ribbon, slightly rotated back view) with location of disease-related mutation indicated (color code in panel). **B-F** Close-up views of some residues outlined in (A).

Table 7 Disease-causing mutations and their effect on hPol III

Subunit	Mutation	Effects on hPol III
Wiedemann-Rautenstrauch syndrome		
RPC1	R1069Q[278]	might destabilize RPC1-RPABC1 interaction and the bridge helix
	D1292N[278]	might destabilize RPC1-RPABC1 interaction
Hypomyelinating leukodystrophy		
RPC1	F558L[280]	might destabilize RPC1-RPC2-RPAC2 interaction
	S602R[331]	might destabilize RPC1 fold intrinsically
	W671R[331]	might destabilize RPC1-RPABC3 interaction
	C724Y[280]	might destabilize RPC1 fold intrinsically
	N775I[280]	might destabilize RPC1-RPC2-RPAC2 interaction
	I804T[333]	might destabilize RPC1-RPC2 interaction and might destabilize the bridge helix
	M852V[280]	might destabilize RPC1-RPC2 interaction and might destabilize the bridge helix
RPC2	L104F[331]	might destabilize RPC2-RPABC4-RPABC5 interaction
	S268G[331]	unclear (neutral effect)
	R442C[331]	might destabilize RPC2-RPABC4-RPABC5 interaction
	T503K[281]	might destabilize RPC2 fold intrinsically
	V523E[281]	might destabilize RPC4/5 interaction/association
	C527R[331]	might destabilize RPC4/5 interaction/association
	R768H[282]	might destabilize RPC2-RPAC1-RPABC5 interaction
	D926E[282]	might destabilize RPC2-RPAC1-RPABC5 interaction
RPAC1	T26I[200,202]	
	T27A[202]	might destabilize interaction platform with RPC1, RPC2, RPAC2 and RPABC3; might
	P30A[202]	destabilize interaction and/or assembly
	N32I[200,202]	
	M65V[200,202]	might destabilize RPAC1-RPAC2 interaction
	N74S[200,202]	might destabilize RPC2-RPAC1-RPAC2 interaction
	V94A[200,202]	unclear (neutral effect)
	I105F[202]	
	H108Y[202]	might destabilize RPC2-RPAC1-RPABC5 interaction
	R109H[200,202]	
	R109C[202]	
	A117P[202]	might destabilize RPAC1 fold intrinsically
	C146R[200,202]	might destabilize RPC2-RPAC1-RPABC5 interaction
	R191Q[200,202]	unclear (solvent exposed)
	I262T[200,202]	might destabilize RPAC1 fold intrinsically
	T313M[202]	might destabilize RPAC1-RPABC5 interaction
	E324K[200,202]	might destabilize RPAC1 fold intrinsically
Treacher-Collins syndrome		
RPAC1	R279Q[199,200]	might destabilize RPAC1 fold intrinsically
	R279W[200]	might destabilize RPAC1 fold intrinsically
RPAC2	E47K[199]	
	T50I[199]	
	L51R[199]	
	G52E[199]	might destabilize RPAC1-RPAC2 interaction and assembly with the core polymerase
	L55V[198]	
	R56C[199]	
	L82S[199]	
	G99S[198,199]	
VZV Susceptibility		
RPC1	R582C[285]	unclear (solvent exposed)
	Q707R[283]	unclear (solvent exposed)

7.2.3 Discussion

Overall, a high degree of conservation concerning functional and structural aspects was found between yeast and human for Pol III. Our results[179] were further validated and complemented by four other independent working groups[191,337–339].

One feature, which was altered in all three human Pols compared to yeast is the foot domain of the largest subunit[222–224,229]. In hPol III, two distinct deletions were recognized within the foot domain, which lead to a rearrangement of the whole domain to a more compact structure. As shown for Pol II, the foot domain is responsible for many protein-protein interactions and a point of enzyme regulation[329,330]. For hPol I, it was proposed that the altered foot domain in humans is a point of protein-protein interactions, as well[224], which suggests that the foot domain in human RNA polymerases add a layer of regulation and additional possibilities for protein-protein interactions.

Pol III subunit RPC10, important for RNA cleavage, backtracking, and termination was found in two distinct conformations concerning the C-ribbon in elongation complexes: inside and outside the funnel[191]. In a backtracked state this domain located inside the funnel, but was less defined due to flexibilities[338], whereas it stayed completely flexible in our monomeric reconstruction[179]. However, in a pre-termination state of human Pol III, the C-ribbon of RPC10 stayed outside the funnel and the poly-dT-termination sequence was trapped in the non-template exit channel[339].

Another difference between yeast and human is the presence of two distinct isoforms of heterotrimer subunit RPC7 (RPC7α and RPC7β). These two isoforms are encoded by two paralogous genes (*POLR3G* and *POLR3GL*), which are developmentally regulated[340]. RPC7α is enriched in cancer and embryonic stem cells[341] and upregulated RPC7α, but not RPC7β, is associated with poor survival outcome for cancer patients[340,342]. One potential explanation for this observation could be that RPC7α seems to be incompatible with binding of the negative Pol III regulator MAF1, as the N-terminus of RPC7α binds to the same region as MAF1[191,337,338].

The conserved C-terminal region of RPC7, present in both isoforms, is essential and can bind the active center cleft, potentially acting in preventing unspecific transcription[337,343]. Additionally, this region would clash with TFIIIB binding and as a consequence might play a role in promoter-dependent Pol III transcription initiation[337].

Another human-specific feature within the built-in transcription factors is the RPC5 C-terminal extension. It includes two tandem winged helix (tWH) domains and plays a role in complete enzyme stability[179]. This metazoan-specific domain, which is not present in lower eukaryotes possesses many phosphorylation sites, which altogether suggests an additional layer of regulation[179].

Having now structures of hPol III in diverse states of the transcription cycle, allowed to map and interpret the consequences of most of the disease-causing mutations. Mutations within six Pol III subunits (RPC1, RPC2, RPAC1, RPAC2, RPC3, RPC6) were identified to cause neurodegenerative disorders in humans. The large number of mutations, the results of mapping and interpreting these mutations, as well as the vast diversity of neurodegenerative phenotypes argued against a single, common disease causing mechanism[313]. A different set of mutations, responsible for an impaired immune response against several viruses, included alterations in four different subunits (RPC1, RPC3, RPC5, RPC6)[313]. This set of mutations seemed to negatively influence predominantly the cytoplasmic activity of Pol III, which is far less understood than the nuclear activity. Nevertheless, a lot of research will be necessary to be able to comprehend the disease-causing mechanisms. The structure and

mapping of mutations represent a valuable tool which will help in small molecule design that are capable of targeting Pol III transcription for therapeutic purposes.
Besides our work, other research groups solved additional reconstructions of hPol III in different states: monomeric hPol III[179,191,337], elongation complexes[191,337,338], during backtracking[338] and in pre-termination state[339]. Similar to RPA12 in hPol I, the C-ribbon of RPC10 is found in two different conformations ('inside' and 'outside' the funnel) in different elongation complexes, but remains outside the funnel in pre-termination state[191,338,339], demonstrating the diverse functions during the transcription cycle of this built-in transcription factor.

The different structures allowed a detailed analysis of structural and functional conservation among species, the comparison across transcription systems, as well as the prediction of the influences disease-causing mutations might have on the enzymes. This information can now support the basic understanding of disease mechanisms as well as drug development and lays the groundwork for further research. While the key concepts of the Pol III activity cycle and its regulatory mechanisms appeared to be conserved, the roles and structural basis of regulation by organism-specific factors in response to external or cell-dependent queues remained poorly understood. Additionally, it remained unclear how the C-terminal acidic tail of subunit RPC7 (C31) may regulate transitions between Pol III transcription cycle phases and for the C-terminal domain of RPC5 further functions besides its importance for whole complex integrity may come to light. Although structures of different states during the transcription cycle could already be solved including elongation, backtracking and pre-termination, phases as initiation, chromatin interaction and in parts termination are still elusive. Especially, recognition of the different promoter types will be challenging to investigate. Specific regulation and re-initiation of Pol III transcription will be another interesting topic of future research. Moreover, cytosolic Pol III activities linked to viral defense may differ from those studied in the nucleus[299], and it will be interesting to investigate which mechanisms might be conserved or diverse comparing nuclear and cytosolic Pol III transcription.

7.3 Comparison of hPol I purification from different cell line systems

Misiaszek *et al.*[222] followed a similar approach as we did by tagging one Pol I subunit endogenously within a human cell line. The difference between the two cell lines was the cell line background and hence, we investigated the potential difference in Pol I yield purifying hPol I from the different cell lines. Two additional cell lines were generated and the yield of hPol I compared to cell culture size or cell pellet weight was tested. One cell line was comparable to the HeLa-POLR1A cell line, but in a HEK293T cell line background as used by Misiaszek *et al.*[222]. For this purpose, the HEK cell line was used as parental cell line and modified via CRISPR/Cas9 with the same protocol and homology directed repair as for our HeLa-POLR1A cell line. The resulting HEK-POLR1A cell line was validated to be homozygous for the sfGFP-tag insertion (see method section 6.5.7). Additionally, an inducible overexpression approach was tested. Here, the T-REx-293 cell line was modified in order to over-express RPA2, RPA34 or RPA43. The overexpressed subunit was again tagged at the C-terminus with sfGFP and a GS-linker in between including a 3C protease cleavage site. The insertion of the tagged subunit sequence was performed as described in method section 6.5.8 and the manufacturer's instructions (Thermo Fisher Scientific) within the desired genomic locus. After validation of the correct insertion, the cells were grown and overexpression of the specifically tagged subunit was induced, respectively.

First of all, cell growth in suspension cell culture of HeLa-POLR1A and HEK-POLR1A was compared, but no major difference was found. Both cell lines grew to a density of roughly $7*10^5$ - $1*10^6$ cells/ml in suspension cell culture and yielded about 1 g of cell pellet for 1 l culture.

In the next step, the cell lines were compared regarding their yield of hPol I in relation to the weight of starting cell material. Here, the exact same single-step purification protocol for hPol I was used without performing an anion-exchange chromatography (see method section 6.7). After the purification, the amount of purified hPol I was compared, but no major difference in yield of hPol I was detected for comparing endogenously tagged hPol I in different cell line backgrounds (HeLa-POLR1A vs. HEK-POLR1A) (Figure 32). For hPol I purification from T-REx cells overexpressing a single subunit (RPA2, RP43, or RPA34 respectively), the purification protocol needs to be optimized as only chromatin was purified. For this purpose, more DNase I and an additional sonification step during cell lysis was established (compare SL1 purification and method section 6.9.2).

With these results, it seemed likely that hPol I amount was well regulated within cells and did not alter to a large extent and as a result, hPol I yield was mostly dependent on the amount of starting cell material and not largely influenced by different cell line background. As a consequence, cell growth and how much cell material can be grown is the major determinant for the amount of *in vitro* purified hPol I.

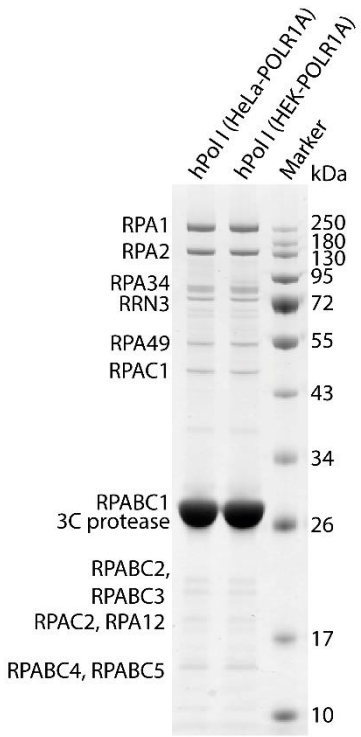


Figure 32 Comparison of hPol I purification from different cell lines

Coomassie-stained SDS-PAGE of purified hPol I from HeLa-POLR1A and HEK-POLR1A cell lines

7.4 hPol I initiation factors

Having structural information about monomeric hPol I and different elongation states[222–224], one of the next steps is to further investigate transcription initiation. Promoter-dependent transcription initiation is tightly regulated and dependent on several transcription factors[10].
The polymerase itself needs to be transferred to an initiation-competent state by binding of RRN3, for which structural information is available[222]. RRN3 and its interaction with hPol I is highly conserved between yeast and human[222]. The hPol I-RRN3 complex is then recruited by the promoter-bound transcription factors SL1 and UBF[14,106]. These promoter-bound factors are not or only partially conserved between yeast and human and many details as well as their structures are elusive[176].
Homodimeric UBF (upstream binding factor) exists in two splice variants, whereas only UBF1, the longer version, binds to the promoter[18,108]. DNA-binding was shown for the first three HMG-boxes after the N-terminal dimerization domain, and consequently the DNA got bended up to a 360° loop as investigated by electron microscopy[17,109,114–116]. The function of HMG-boxes 4-6 is rather unclear and proposed to mainly function in protein-protein interactions[176]. The acidic C-terminal domain (CTD) of UBF functions in protein-protein interaction and interacts with SL1 via TAF1A and TBP[11,17,117–119]. Complete structural information is lacking for UBF, but most of the HMG-boxes alone were subject to structural investigations. NMR solution structures are available for HMG-box 1 (human, PDB 1K99), 3 (mouse, PDB 1V64), 4 (mouse, PDB 1WGF), 5 (human, PDB 1L8Y/1L8Z) and 6 (mouse, PDB 1V63) as well as an X-ray structure of human HMG-box 5 (PDB 2HDZ)[14,344,345], whereas the amino acid sequences of mouse and human UBF is 98 % identical. Although all these structures together with the low resolution information from electron microscopy of *Xenopus laevis* xUBF bound to DNA[346] already created an overall frame, the detailed mechanisms of DNA-binding, promoter recognition and its role in transcription initiation and structural features are still elusive.
Multi-subunit factor SL1 (selectivity factor 1) binds cooperatively with UBF to the promoter DNA and is responsible for species-specificity[17,70,100,105,108]. In parts, subunits TAF1A, TAF1B, and TAF1C are homologues to yeast core factor subunits Rrn11, Rrn7, and Rrn6, respectively. Additionally, subunits TAF1D and TBP are stably associated in SL1[17,70,99,100], but structural information is missing for the complex besides TBP alone[347,348]. Moreover, little is known about the cooperativity with UBF in promoter-binding and the recruitment of the hPol I-RRN3 complex.
For structural and functional investigations of the three transcription initiation factors RRN3, SL1, and UBF, *in vitro* purification protocols needed to be established. For this purpose, a reasonable expression system, which allows proper folding and supports a stable structure, needed to be found first, before establishing and optimizing a purification protocol with sufficient buffers.

For RRN3 expression, insect cell (SF21) culture was chosen, but later on it was shown, that RRN3 expression was also possible in the *E. coli* background[222] and expression system was changed (see method section 6.9.1).
For expression in SF21 cells, N-terminal His-MBP-tagged (3C-cleavable) RRN3 was cloned into bacmids to be able to produce virus in SF21 cells. For expression, a large number of SF21 cells were infected with this virus and expression was performed for 2.5 days. After harvesting the cells, cell lysis and an MBP-pulldown was performed. After elution of the tagged RRN3 with maltose, a size exclusion

chromatography was performed either directly or after 3C cleavage of the tag. Fractions with monomeric RRN3 were collected and the purification was analyzed via SDS-PAGE (Figure 33A). This purification protocol of RRN3, recombinantly expressed in SF21 cells, yielded only low amounts of monomeric RRN3, whereas a large fraction of RRN3 was aggregating or formed multimers.
As it was shown that human RRN3 can also be overexpressed in *E. coli*, a comparable strategy to the published protocol was tested[222]. After cell lysis, a His-pulldown was performed to trap N-terminal His-tagged RRN3 and eluted with imidazole, before a HiTrap Q chromatography was conducted. In the last step, a size exclusion chromatography was executed and the RRN3 containing fractions were pooled and concentrated (Figure 33A).

For the multi-subunit factor SL1, a human expression system was chosen. Therefore, the cell line T-REx-293 was modified, so that the overexpression of a C-terminal sfGFP-tagged (3C cleavable) TAF1A version could be induced. After two days of overexpressing TAF1A, the cells were harvested. With the cell lysate a GFP-pulldown and elution with 3C protease was performed similar to the hPol I purification (see method section 6.9.2). SDS-PAGE and mass spectrometry confirmed that the five subunits TAF1A, TAF1B, TAF1C, TAF1D, and TBP were present in the elution, but TAF12, a subunit which was proposed to be part of SL1 was not detectable[101] (Figure 33B).

The two splice variants of UBF were cloned from cDNA and a 3C-cleavable N-terminal His-tag was used. Even though the DNA-binding factor was toxic for bacteria in the longer term, expression conditions were established in BL21 cells, after getting advice from Prof. Tom Moss (see method section 6.9.3). After cell lysis, a His-pulldown was performed including an ATP-wash step to get rid of chaperones and bound DNA. After elution, a HiTrap Heparin column was executed, before the final size exclusion chromatography. UBF-containing fractions were pooled and concentrated (Figure 33C). The molecular weight was measured with a mass photometer, which verified the homodimerization. Preliminary results of electro-mobility shift assays showed that DNA-binding was possible for the *in vitro* purified UBF1.

Having now purification protocols for all three human transcription initiation factors allowed to start with systematic structure-function analysis *in vitro*, including analysis of DNA-binding and preferred sequences, of protein-protein interactions and structural analysis of the factors alone, in a DNA-bound state and in a complex.

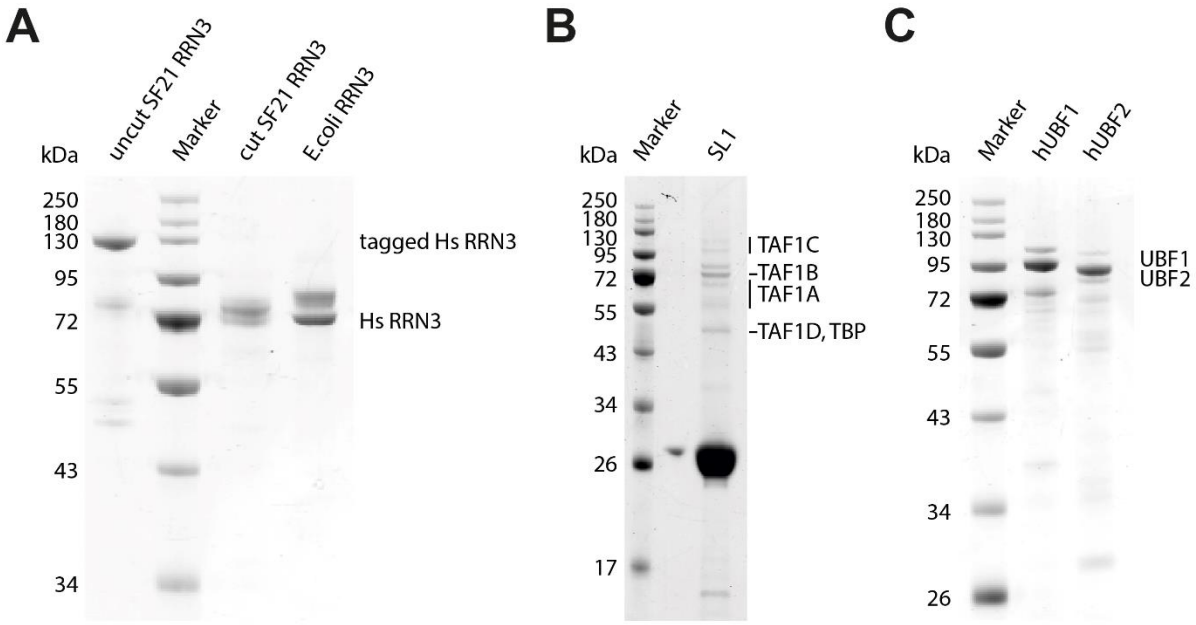


Figure 33 Purification of human Pol I transcription factors RRN3, SL1 and UBF1 and UBF2

A-C Coomassie-stained SDS-PAGE of *in vitro* purified human Pol I transcription initiation factors RRN3 (A), SL1 (B) and UBF (C).

8 List of Figures

9 List of Tables

10 Abbreviations

A	Ampere
Å	Angstrom
aa	amino acid
AML	acute myeloid leukemia
AMPK	AMP-activated protein kinase
ARM	charged tail region of RPA34
ATP	Adenosintriphophate
Bdp1	B double prime 1
bp	base pair
Brf1/2	B-related factor 1/2
BSA	bovine serum albumin
Cas9n	CRISPR associated protein 9 nickase
CBP	CREB-binding protein
CF	core factor
ChIP	Chromatin Immuno-precipitation
CK2	casein kinase 2
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
cryo-EM	electron cryo-microscopy
C-ribbon	C-terminal ribbon
C-terminus	Carboxy-terminus (of proteins)
CTD	C-terminal domain
CTP	Cytosintriphophate
CTR	C-terminal region
Da	Dalton
DAPI	4’,6-diamidino-2-phenylindole
DMEM	Dulbeccos Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleotide acid
dNTP	desoxyribonucleosidtriphosphate
ds	double-strand
DSB	double-strand break
EC	elongation complex
EM	electron microscopy
EMSA	electrophoretic mobility shift assay
ERK	extracellular signal-regulated kinase
ETS	external transcribed spacer
FACS	fluorescence-activated cell sorting’
FACT	facilitates chromatin transcription
FBS	fetal bovine serum

FL	full-length
g	gramm / * g: gravitational acceleration
gDNA	genomic DNA
GFP	green fluorescent protein
GO	graphene oxide
gRNA	guide RNA
GS-linker	glycine-serine linker
GTP	Guanosintriphophate
h	hour
HDR	homology-directed repair
HeLa	cell line derivated from Henrietta Lacks
HEK	human embryonic kidney cell line
His	Histidine
HL	Hypomyelinating Leukodystrophy
HMG	high mobility group
hPol	human RNA polymerase
HRV 3C	Human Rhinovirus (HRV) 3C Protease
Hs (H. sapiens)	Homo sapiens
HTH	helix-turn-helix
ICR	internal control region
IF	immunofluorescence
IgG	Immunoglobulin G
IGS	intergenic spacer
IPTG	Isopropyl β-d-1-thiogalactopyranoside
ITS	internal transcribed spacer
JNK2	Jun N-terminal kinase 2
kb	kilobase
keV	kilo electronvolt
L	liter
LB	Lysogeny broth
LoG	Laplacian-of-Gaussian
M	mol/L (molar concentration)
MAPK	mitogen-activated protein kinase
MBP	maltose-binding protein
min	minute
miRNA	microRNA
mol	mole (amount of substance)
MOPS	3-morpholinopropanesulfonic acid
mRNA	messenger RNA
mTOR	mammalian target of rapamycin
NDR	nucleosome depleted region
NHEJ	non-homologous end junction
Ni-NTA	nickel-nitrilotriacetic acid

NOR	nucleolar organizer region
N-ribbon	N-terminal ribbon
nt	nucleotide or non-transcribed/no-template
N-terminus	Amino-terminus (of proteins)
NTP	nucleotidetriphosphate
OC	open complex
OD	optical density
ORF	open reading frame
PAM	protospacer adjacent motif
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEI	polyethylenimine
PIC	pre-initiation complex
PKC	protein kinase C
Pol	DNA-dependent RNA polymerase
PP2A	phosphatase type 2A
pre-rRNA	precursor ribosomal RNA
PSE	proximal sequence element
PTM	post-transcriptional modifications
PTRF	Pol I transcript release factor
PVDF	polyvinylidene difluoride
rDNA	ribosomal DNA
RFB	replication fork barrier
RNA	ribonucleotide acid
RPC5EXT	C-terminal extension of Pol III subunit RPC5
rpm	rounds per minute
r-protein	ribosomal protein
rRNA	ribosomal RNA
RSK	ribosomal S6 kinase
s	second
Sc (*S. cerevisiae*)	*Saccharomyces cerevisiae*
SD	standard deviation
SDS	Sodium dodecyl sulfate
SEC	size exclusion chromatography
sfGFP	superfolder green fluorescent protein
SF21	*Spodoptera frugiperda* 21 insect cell line
SIRT7	NAD-dependent protein deacetylase sirtuin-7
SL1	selectivity factor 1
SNAPc	snRNA activating protein complex
SOB	Super Optimal Broth
Sp (*S. pombe*)	*Schizosaccharomyces pombe*
ss	single-strand

TAF	TATA-Box Binding Protein Associated Factor
TB	Terrific Broth
TBE	Tris/Borate/EDTA
TBP	TATA-box-binding protein
TCS	Treacher Collin Syndrome
T-Rex-293	Tetracycline-Regulated Expression HEK293T cell line
tRNA	transfer RNA
TSS	transcription start site
TTF-I	transcription termination factor I
tWH	tandem winged helix
U	unit
UAF	upstream activating factor
UBF	upstream binding factor
UCE	upstream control element
UPE	upstream promoter element
UTP	uridine-5’-triphosphate
UTR	untranslated region
V	Volt
VZV	Varicella zoster virus
v/v	volume/volume (volume fraction)
W	Watt
WB	western blot
WRS	Wiedemann-Rautenstrauch syndrome
w/v	weight/volume (mass concentration)
°C	degree celcius
%	percent

11 Acknowledgements

The last four years I had the pleasure to meet and get to know many wonderful and inspiring characters while working on this project. Without any of them, it wouldn't be the same. Thanks a lot to everybody involved!

Special and enormous thanks go to Prof. Dr. Christoph Engel, my supervisor and father of this project. Christoph trusted and believed in me from the start on, committing me this exciting and challenging project about human Pol I. Right from the beginning, he engaged himself with long sessions at the electron microscope and taught me data processing to get me started. He has always an open ear and supports if problems arise. He is also encouraging me as well as the whole team to self-confidence and pursues independent thinking by thoughtful guidance, asking the right questions, and believing in one's own knowledge and finding the solution by yourself. I'm still lucky having found this place and I'm looking forward to explore further this interesting research field.

My colleague Dr. Michael Pilsl proved to be a reliable fountain of scientific knowledge. Michi is the master of the new cryoARM200, a fabulous OddPols-buddy in Spain 2023, and the only one sharing my passion for soccer. Discussing soccer news was sometimes a good relief intercepting deadlocked thoughts. Philipp Becker and Florian Heiß, my co-PhD students who already left the lab, introduced me to the lab back then and we shared countless hours there. Flo always described himself as "problem-solver", taking care of all practical, hands-on issues. He was a great source for discussions about the project and left a gap in the lab with his calm, balanced, and communicative personality. Philipp's presence was always pleasant and grounded the atmosphere in the room. "His" method, the crystallization, enriched our lab's expertise and he taught me everything, I know about it so far. I only got to know Lena Heuschneider, our former lab technician, for a couple of weeks, but I appreciated her a lot and I am very thankful for her growing human cells before I started and introducing me to the cultivation of human suspension cells. Mona Höcherl, our lab technician, purified a lot of recombinant proteins for me, and assisted during cloning. Mona is our unofficial head of lab procedures, maintains all our equipment, keeps an eye on all commonly used goods, organizes almost everything, and last but not least supplies the whole group with fresh stone-oven bread. Mona, Michi, Flo, and Philipp: thanks for accompanying me through this thesis, for the fruitful discussions, and all the cheerful (lunch) breaks!

Not going unmentioned should be Matthias Breindl, our latest PhD student, and Stefan Jungwirth, our Master student, who have been good co-workers in our lab team - Matthias bringing back Philipp's taste for music and Stefan having been a pleasant, eager and hard-working student. All "my" students - Lea, Eva, Anna, and Patrick - performed their tasks in a very responsible manner which truly contributed to the success of my project. Lea worked on nucleosome passaging of yeast Pol I and helped during cloning, whereas Eva, Anna, and Patrick put great effort into optimizing the purifications of the two transcription factors SL1 and UBF.

Gisela Pöll, being officially part of the Biochemistry III department, feels like a direct colleague of us, answers all questions, supports everyone, and is often even without speaking just through her presence an inspiration and the solution of my then unasked questions. Our two engineers Norbert Eichner and Gerhard Lehmann from the Biochemistry I department have been doing a great job in setting-up

and maintaining all our computational infrastructure as well as advising on and helping with all our technical jumble. I'm owing you all!

Many thanks go also to RiGEL and especially to Kinga Ay and Carolin Apfel for all their efforts, organizing all the brilliant method or carrier-supporting courses, summer schools, get-together events, and their willingness to help with any kind of paperwork. Additionally, they offered via RiGEL the opportunity to go to London for a research stay at the Institute of Cancer Research to learn a lot from my collaborators Prof. Dr. Alessandro Vannini, Dr. Ewan Ramsay, Dr. Guillermo Abascal-Palacios during my first year.
This collaboration persisted throughout my whole PhD time, started with the wet-lab work together in London, followed by fruitful discussions and protocol exchanges, and leading to great and continuous cooperation in different aspects. I want to express my gratitude to you all for your patience with me and hope to carry on our enriching collaboration.

Prof. Dr. Dina Grohmann and Prof. Dr. Carrie Bernecky functioned as my two mentors and hereby acted as inspiring role models. It was a great pleasure to have them at my side during this exciting time. They always supported me in my research, gave good advice, and offered help whenever needed, leading to awesome collaborations with each of them.
Prof. Dr. Carrie Bernecky and Dr. Katarina Tlučková hosted me at the Institute of Science and Technology Austria even during the hard times of Corona to teach me how to prepare graphene-oxide supported cryo-EM grids. Moreover, they offered the opportunity to screen and collect data at their microscopes, which led to the breakthrough in cryo-EM sample optimization. I'm much obliged to your efforts, thanks a lot!
Prof. Dr. Dina Grohmann and Andreas Schmidbauer, together with Christoph and me established an unofficial "human Pol I research team" at the University of Regensburg. The approach from two different angles – the analysis of single-molecule dynamics and the structure-function analysis – inspires the ideas and discussions and therefore improves the outcoming results. There is an active exchange of knowledge, materials, and ideas and I really love the open discussions we have. Thank you all!

At the University of Regensburg, there are several research groups, I want to mention for their help. Prof. Dr. Remco Sprangers allowed us to install a human cell culture at his department and he himself, Johanna Stöfl and his whole team are always very helpful and are supporting the routine maintenance. Elisabeth Silberhorn and Helen Hoffmeister are both colleagues, I worked with from the beginning in the human cell culture. Later on, they contributed a lot of knowledge about how to install a cell culture and are a great source for advice. Dr. Astrid Bruckmann and her group performed mass spectrometry and supported with data analysis and evaluation. Dr. Kristina Straub executed all the phylogenetic analysis and had always time to discuss upcoming problems. Dr. Andrea Bleckmann introduced me to confocal microscopy, taught me everything I know about it and was always willing to advise. Prof. Dr. Reinhard Rachel taught me a lot about electron microscopy and spend many hours to introduce me to the JEM-2100F. Dr. Philipp Milkereit enlarged my knowledge about electron microscopy even further and is always open for discussions. Prof. Dr. Joachim Griesenbeck proved to be a very reliable contact person with a profound knowledge combined with his own sympathetic way to communicate. Prof. Dr. Herbert Tschochner is always open for discussion and especially for *in vitro* transcriptions of Pol I a great advisor. He had also enabled my inspiring stay at the University of Colorado, Boulder with Prof. Dr. Tom Cech. Tom and his team taught me a lot and he always believed in my strengths, guided me

to a self-reliant worker and is still a mentor of mine, thank you. Prof. Dr. Gunter Meister was already an important person during my studies, supervising my Bachelor and Master thesis, and thereby setting a profound basic education. He and his team have been giving me valuable advice in designing CRISPR/Cas9 genome-editing, establishing T-REx cell lines, and performing siPool knock-downs. The good working atmosphere, generated by all the actual and former colleagues working at the Structural Biochemistry, Biochemistry I and III departments, have been leading to many good scientific and private conversations and consequently, I enjoyed going to work every day. I'm very grateful for all their time and their best efforts they invested in myself and my research.
Additional to all the wonderful persons mentioned above, I need to thank all our passionate wide-spread collaborators, who participated in the success of this project. Dr. Philip Gunkel and Prof. Dr. Volker Cordes provided the HeLa P2 and HeLa-POLR1C cell lines and gave good advice on designing CRISPR/Cas9 genome-editing experiments. Prof. Dr. Tom Moss has been a great source for fresh ideas and new viewpoints, especially during his stays in Regensburg with elongated scientific discussions. Together with Dr. Jean-Clément Mars, he performed the ChIP-Seq analysis. Prof. Dr. Till Rudack and Torben Fürtges executed molecular modelling of Top2a interactions. Prof. Dr. Valerie Lamour provided purified Top2a full-length and ΔCTD and gave good advice for handling. Dr. Konstantin Panov and Dr. Joost Zomerdijk purified recombinant fUBF and performed Top2a co-IPs and fUBF-Top2a pulldowns.

I want to thank all the members of my doctoral committee Prof. Dr. Christoph Engel, Prof. Dr. Dina Grohmann, Prof. Dr. Herbert Tschochner and Prof. Dr. Klaus Grasser to function as referees and examiners as well as Prof. Dr. Gunter Meister for chairing the thesis defense and taking their time.

Last but not least, I want to thank my family and friends for their support! Vroni, a former colleague and friend for a long time, has been a good balance and support in demanding times. My brother Fabi is always lifting my spirits with his open and positive character and shows me often the ease of being - *just do it*. Joy, being the best companion, positively influenced and formed the best of me. I learned a lot from you. My parents, Susi and Armin, always lovingly support me. They taught me curiosity in and for life, and it's always fun coming home and also enjoying the outdoors together with Pearly. After all, I need to thank Flo: you not only understand the demands of occupying science (or soccer), but you are also always there for me, bringing me down to earth, treating me (with fabulous cooked meals) and I love to explore the world with you.

12 References

1. Crick, F. Central dogma of molecular biology. *Nature* **227,** 561–563; 10.1038/227561a0 (1970).
2. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224,** 234–237; 10.1038/224234a0 (1969).
3. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nature reviews. Microbiology* **9,** 85–98; 10.1038/nrmicro2507. (2011).
4. Werner, F. Structural evolution of multisubunit RNA polymerases. *Trends in microbiology* **16,** 247–250; 10.1016/j.tim.2008.03.008 (2008).
5. Engel, C., Neyer, S. & Cramer, P. Distinct Mechanisms of Transcription Initiation by RNA Polymerases I and II. *Annual review of biophysics* **47,** 425–446; 10.1146/annurev-biophys-070317-033058 (2018).
6. Girbig, M., Misiaszek, A. D. & Müller, C. W. Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nature reviews. Molecular cell biology* **23,** 603–622; 10.1038/s41580-022-00476-9 (2022).
7. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Molecular cell* **45,** 439–446; 10.1016/j.molcel.2012.01.023 (2012).
8. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nature reviews. Molecular cell biology* **16,** 129–143; 10.1038/nrm3952. (2015).
9. Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M. & Pagano, A. The expanding RNA polymerase III transcriptome. *Trends in genetics : TIG* **23,** 614–622; 10.1016/j.tig.2007.09.001. (2007).
10. Goodfellow, S. J. & Zomerdijk, J. C. B. M. Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes. *Sub-cellular biochemistry* **61,** 211–236; 10.1007/978-94-007-4525-4_10 (2013).
11. Moss, T., Langlois, F., Gagnon-Kugler, T. & Stefanovsky, V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cellular and molecular life sciences : CMLS* **64,** 29–49; 10.1007/s00018-006-6278-1 (2007).
12. Pitts, S. & Laiho, M. Regulation of RNA Polymerase I Stability and Function. *Cancers* **14**; 10.3390/cancers14235776 (2022).
13. Klinge, S. & Woolford, J. L. Ribosome assembly coming into focus. *Nature reviews. Molecular cell biology* **20,** 116–131; 10.1038/s41580-018-0078-y (2019).
14. Rong, H. *et al.* Structure of human upstream binding factor HMG box 5 and site for binding of the cell-cycle regulatory factor TAF1. *Acta crystallographica. Section D, Biological crystallography* **63,** 730–737; 10.1107/S0907444907017027 (2007).
15. Grummt, I. Regulation of mammalian ribosomal gene transcription by RNA polymerase I. *Progress in nucleic acid research and molecular biology* **62,** 109–154; 10.1016/s0079-6603(08)60506-1 (1999).
16. Reeder, R. H. Regulation of RNA polymerase I transcription in yeast and vertebrates. *Progress in nucleic acid research and molecular biology* **62,** 293–327; 10.1016/s0079-6603(08)60511-5 (1999).
17. Russell, J. & Zomerdijk, J. C. B. M. The RNA polymerase I transcription machinery. *Biochemical Society symposium,* 203–216; 10.1042/bss0730203 (2006).
18. Tremblay, M. G. *et al.* Ribosomal DNA promoter recognition is determined in vivo by cooperation between UBTF1 and SL1 and is compromised in the UBTF-E210K neuroregression syndrome. *PLoS genetics* **18,** e1009644; 10.1371/journal.pgen.1009644 (2022).
19. Herdman, C. *et al.* A unique enhancer boundary complex on the mouse ribosomal RNA genes persists after loss of Rrn3 or UBF and the inactivation of RNA polymerase I transcription. *PLoS genetics* **13,** e1006899; 10.1371/journal.pgen.1006899 (2017).

20. Moss, T. At the crossroads of growth control; making ribosomal RNA. *Current opinion in genetics & development* **14,** 210–217; 10.1016/j.gde.2004.02.005 (2004).

21. Jorgensen, P. *et al.* A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes & development* **18,** 2491–2505; 10.1101/gad.1228804 (2004).

22. Vannini, A. A structural perspective on RNA polymerase I and RNA polymerase III transcription machineries. *Biochimica et biophysica acta* **1829,** 258–264; 10.1016/j.bbagrm.2012.09.009. (2013).

23. McStay, B. Nucleolar organizer regions: genomic 'dark matter' requiring illumination. *Genes & development* **30,** 1598–1610; 10.1101/gad.283838.116 (2016).

24. Schöfer, C. & Weipoltshammer, K. Nucleolus and chromatin. *Histochemistry and cell biology* **150,** 209–225; 10.1007/s00418-018-1696-3 (2018).

25. Henderson, A. S., Warburton, D. & Atwood, K. C. Location of ribosomal DNA in the human chromosome complement. *Proceedings of the National Academy of Sciences of the United States of America* **69,** 3394–3398; 10.1073/pnas.69.11.3394. (1972).

26. Roussel, P., André, C., Comai, L. & Hernandez-Verdun, D. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *The Journal of cell biology* **133,** 235–246; 10.1083/jcb.133.2.235 (1996).

27. Cheutin, T. *et al.* Three-dimensional organization of active rRNA genes within the nucleolus. *Journal of cell science* **115,** 3297–3307; 10.1242/jcs.115.16.3297 (2002).

28. Scheer, U., Thiry, M. & Goessens, G. Structure, function and assembly of the nucleolus. *Trends in cell biology* **3,** 236–241; 10.1016/0962-8924(93)90123-i (1993).

29. Thompson, M., Haeusler, R. A., Good, P. D. & Engelke, D. R. Nucleolar clustering of dispersed tRNA genes. *Science (New York, N.Y.)* **302,** 1399–1401; 10.1126/science.1089814 (2003).

30. Matera, A. G., Frey, M. R., Margelot, K. & Wolin, S. L. A perinucleolar compartment contains several RNA polymerase III transcripts as well as the polypyrimidine tract-binding protein, hnRNP I. *The Journal of cell biology* **129,** 1181–1193; 10.1083/jcb.129.5.1181 (1995).

31. Németh, A. *et al.* Initial genomics of the human nucleolus. *PLoS genetics* **6,** e1000889; 10.1371/journal.pgen.1000889 (2010).

32. Zomerdijk, J. Structural biology: Pivotal findings for a transcription machine. *Nature* **502,** 629–630; 10.1038/nature12700 (2013).

33. Miller, O. L. & Beatty, B. R. Visualization of nucleolar genes. *Science (New York, N.Y.)* **164,** 955–957; 10.1126/science.164.3882.955 (1969).

34. Miller, O. L. & Bakken, A. H. Morphological studies of transcription. *Acta endocrinologica. Supplementum* **168,** 155–177; 10.1530/acta.0.071s155 (1972).

35. Osheim, Y. N. *et al.* Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae. *Molecular cell* **16,** 943–954; 10.1016/j.molcel.2004.11.031 (2004).

36. Neyer, S. *et al.* Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* **540,** 607–610; 10.1038/nature20561 (2016).

37. Potapova, T. A. & Gerton, J. L. Ribosomal DNA and the nucleolus in the context of genome organization. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **27,** 109–127; 10.1007/s10577-018-9600-5 (2019).

38. Jones, M. H., Learned, R. M. & Tjian, R. Analysis of clustered point mutations in the human ribosomal RNA gene promoter by transient expression in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **85,** 669–673; 10.1073/pnas.85.3.669 (1988).

39. Engel, C. *et al.* Structural Basis of RNA Polymerase I Transcription Initiation. *Cell* **169,** 120-131.e22; 10.1016/j.cell.2017.03.003 (2017).

40. Cramer, P. *et al.* Architecture of RNA polymerase II and implications for the transcription mechanism. *Science (New York, N.Y.)* **288,** 640–649; 10.1126/science.288.5466.640 (2000).

41. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science (New York, N.Y.)* **292,** 1863–1876; 10.1126/science.1059493 (2001).

42. Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502,** 650–655; 10.1038/nature12712 (2013).

43. Fernández-Tornero, C. *et al.* Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502,** 644–649; 10.1038/nature12636 (2013).

44. Heiss, F. B., Daiß, J. L., Becker, P. & Engel, C. Conserved strategies of RNA polymerase I hibernation and activation. *Nature communications* **12,** 758; 10.1038/s41467-021-21031-8 (2021).

45. Fernández-Tornero, C. RNA polymerase I activation and hibernation: unique mechanisms for unique genes. *Transcription* **9,** 248–254; 10.1080/21541264.2017.1416267 (2018).

46. Weinzierl, R. O. J. The Bridge Helix of RNA polymerase acts as a central nanomechanical switchboard for coordinating catalysis and substrate movement. *Archaea (Vancouver, B.C.)* **2011,** 608385; 10.1155/2011/608385 (2011).

47. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nature structural & molecular biology* **15,** 811–818; 10.1038/nsmb.1458 (2008).

48. Brueckner, F. *et al.* Structure-function studies of the RNA polymerase II elongation complex. *Acta crystallographica. Section D, Biological crystallography* **65,** 112–120; 10.1107/S0907444908039875 (2009).

49. Torreira, E. *et al.* The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription. *eLife* **6**; 10.7554/eLife.20832 (2017).

50. Tafur, L. *et al.* Molecular Structures of Transcribing RNA Polymerase I. *Molecular cell* **64,** 1135–1143; 10.1016/j.molcel.2016.11.013 (2016).

51. Schwank, K. *et al.* RNA polymerase I (Pol I) lobe-binding subunit Rpa12.2 promotes RNA cleavage and proofreading. *Journal of Biological Chemistry* **298,** 101862; 10.1016/j.jbc.2022.101862 (2022).

52. Jennebach, S., Herzog, F., Aebersold, R. & Cramer, P. Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. *Nucleic acids research* **40,** 5591–5601; 10.1093/nar/gks220 (2012).

53. Kettenberger, H., Armache, K.-J. & Cramer, P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114,** 347–357; 10.1016/s0092-8674(03)00598-1 (2003).

54. Scull, C. E., Lucius, A. L. & Schneider, D. A. The N-terminal domain of the A12.2 subunit stimulates RNA polymerase I transcription elongation. *Biophysical journal* **120,** 1883–1893; 10.1016/j.bpj.2021.03.007 (2021).

55. Tafur, L. *et al.* The cryo-EM structure of a 12-subunit variant of RNA polymerase I reveals dissociation of the A49-A34.5 heterodimer and rearrangement of subunit A12.2. *eLife* **8**; 10.7554/eLife.43204 (2019).

56. Geiger, S. R. *et al.* RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Molecular cell* **39,** 583–594; 10.1016/j.molcel.2010.07.028 (2010).

57. Kuhn, C.-D. *et al.* Functional architecture of RNA polymerase I. *Cell* **131,** 1260–1272; 10.1016/j.cell.2007.10.051 (2007).

58. van Mullem, V., Landrieux, E., Vandenhaute, J. & Thuriaux, P. Rpa12p, a conserved RNA polymerase I subunit with two functional domains. *Molecular microbiology* **43,** 1105–1113; 10.1046/j.1365-2958.2002.02824.x (2002).

59. Beckouet, F. *et al.* Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription. *Molecular and cellular biology* **28,** 1596–1605; 10.1128/MCB.01464-07 (2008).

60. McNamar, R., Rothblum, K. & Rothblum, L. I. The Mammalian and Yeast A49 and A34 Heterodimers: Homologous but Not the Same. *Genes* **12**; 10.3390/genes12050620 (2021).

61. Pilsl, M. *et al.* Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nature communications* **7,** 12126; 10.1038/ncomms12126 (2016).

62. Han, Y. *et al.* Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *eLife* **6**; 10.7554/eLife.27414 (2017).

63. Sadian, Y. *et al.* Molecular insight into RNA polymerase I promoter recognition and promoter melting. *Nature communications* **10,** 5543; 10.1038/s41467-019-13510-w (2019).

64. Merkl, P. E. *et al.* RNA polymerase I (Pol I) passage through nucleosomes depends on Pol I subunits binding its lobe structure. *The Journal of biological chemistry* **295,** 4782–4795; 10.1074/jbc.RA119.011827 (2020).

65. Albert, B. *et al.* RNA polymerase I-specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle. *The Journal of cell biology* **192,** 277–293; 10.1083/jcb.201006040 (2011).

66. Penrod, Y., Rothblum, K. & Rothblum, L. I. Characterization of the interactions of mammalian RNA polymerase I associated proteins PAF53 and PAF49. *Biochemistry* **51,** 6519–6526; 10.1021/bi300408q (2012).

67. Chen, S. *et al.* Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Molecular cell* **52,** 303–313; 10.1016/j.molcel.2013.10.010 (2013).

68. Hannan, R. D. *et al.* Affinity purification of mammalian RNA polymerase I. Identification of an associated kinase. *The Journal of biological chemistry* **273,** 1257–1267; 10.1074/jbc.273.2.1257 (1998).

69. Moorefield, B., Greene, E. A. & Reeder, R. H. RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 4724–4729; 10.1073/pnas.080063997 (2000).

70. Comai, L., Tanese, N. & Tjian, R. The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. *Cell* **68,** 965–976; 10.1016/0092-8674(92)90039-f (1992).

71. Knutson, B. A. & Hahn, S. Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. *Science (New York, N.Y.)* **333,** 1637–1640; 10.1126/science.1207699 (2011).

72. Knutson, B. A., Luo, J., Ranish, J. & Hahn, S. Architecture of the Saccharomyces cerevisiae RNA polymerase I Core Factor complex. *Nature structural & molecular biology* **21,** 810–816; 10.1038/nsmb.2873 (2014).

73. Knutson, B. A. & Hahn, S. TFIIB-related factors in RNA polymerase I transcription. *Biochimica et biophysica acta* **1829,** 265–273; 10.1016/j.bbagrm.2012.08.003 (2013).

74. Keener, J., Dodd, J. A., Lalo, D. & Nomura, M. Histones H3 and H4 are components of upstream activation factor required for the high-level transcription of yeast rDNA by RNA polymerase I. *Proceedings of the National Academy of Sciences of the United States of America* **94,** 13458–13462; 10.1073/pnas.94.25.13458 (1997).

75. Baudin, F. *et al.* Mechanism of RNA polymerase I selection by transcription factor UAF. *Science advances* **8,** eabn5725; 10.1126/sciadv.abn5725 (2022).

76. Hall, D. B., Wade, J. T. & Struhl, K. An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in Saccharomyces cerevisiae. *Molecular and cellular biology* **26,** 3672–3679; 10.1128/MCB.26.9.3672-3679.2006 (2006).

77. Merz, K. *et al.* Actively transcribed rRNA genes in S. cerevisiae are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. *Genes & development* **22,** 1190–1204; 10.1101/gad.466908 (2008).

78. Wittner, M. *et al.* Establishment and maintenance of alternative chromatin states at a multicopy gene locus. *Cell* **145,** 543–554; 10.1016/j.cell.2011.03.051 (2011).

79. Albert, B. *et al.* Structure-function analysis of Hmo1 unveils an ancestral organization of HMG-Box factors involved in ribosomal DNA transcription from yeast to human. *Nucleic acids research* **41,** 10135–10149; 10.1093/nar/gkt770 (2013).

80. Schächner, C. *et al.* Establishment and Maintenance of Open Ribosomal RNA Gene Chromatin States in Eukaryotes. *Methods in molecular biology (Clifton, N.J.)* **2533,** 25–38; 10.1007/978-1-0716-2501-9_2 (2022).

81. Hannig, K. *et al.* The C-terminal region of Net1 is an activator of RNA polymerase I transcription with conserved features from yeast to human. *PLoS genetics* **15,** e1008006; 10.1371/journal.pgen.1008006 (2019).

82. Warner, J. R. The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* **24,** 437–440; 10.1016/s0968-0004(99)01460-7 (1999).

83. Smith, M. L., Cui, W., Jackobel, A. J., Walker-Kopp, N. & Knutson, B. A. Reconstitution of RNA Polymerase I Upstream Activating Factor and the Roles of Histones H3 and H4 in Complex Assembly. *Journal of molecular biology* **430,** 641–654; 10.1016/j.jmb.2018.01.003. (2018).

84. Oakes, M., Siddiqi, I., Vu, L., Aris, J. & Nomura, M. Transcription factor UAF, expansion and contraction of ribosomal DNA (rDNA) repeats, and RNA polymerase switch in transcription of yeast rDNA. *Molecular and cellular biology* **19,** 8559–8569; 10.1128/MCB.19.12.8559. (1999).

85. Keener, J., Josaitis, C. A., Dodd, J. A. & Nomura, M. Reconstitution of yeast RNA polymerase I transcription in vitro from purified components. TATA-binding protein is not required for basal transcription. *The Journal of biological chemistry* **273,** 33795–33802; 10.1074/jbc.273.50.33795. (1998).

86. Steffan, J. S., Keys, D. A., Dodd, J. A. & Nomura, M. The role of TBP in rDNA transcription by RNA polymerase I in Saccharomyces cerevisiae: TBP is required for upstream activation factor-dependent recruitment of core factor. *Genes & development* **10,** 2551–2563; 10.1101/gad.10.20.2551. (1996).

87. Kramm, K., Engel, C. & Grohmann, D. Transcription initiation factor TBP: old friend new questions. *Biochemical Society transactions* **47,** 411–423; 10.1042/BST20180623 (2019).

88. Bric, A., Radebaugh, C. A. & Paule, M. R. Photocross-linking of the RNA polymerase I preinitiation and immediate postinitiation complexes: implications for promoter recruitment. *The Journal of biological chemistry* **279,** 31259–31267; 10.1074/jbc.M311828200 (2004).

89. Blattner, C. *et al.* Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth. *Genes & development* **25,** 2093–2105; 10.1101/gad.17363311 (2011).

90. Engel, C., Plitzko, J. & Cramer, P. RNA polymerase I-Rrn3 complex at 4.8 Å resolution. *Nature communications* **7,** 12129; 10.1038/ncomms12129 (2016).

91. Bodem, J. *et al.* TIF-IA, the factor mediating growth-dependent control of ribosomal RNA synthesis, is the mammalian homolog of yeast Rrn3p. *EMBO reports* **1,** 171–175; 10.1093/embo-reports/kvd032 (2000).

92. Cavanaugh, A. H. *et al.* Rrn3 phosphorylation is a regulatory checkpoint for ribosome biogenesis. *The Journal of biological chemistry* **277,** 27423–27432; 10.1074/jbc.M201232200 (2002).

93. Grummt, I. & Voit, R. Linking rDNA transcription to the cellular energy supply. *Cell cycle (Georgetown, Tex.)* **9,** 225–226; 10.4161/cc.9.2.10614 (2010).

94. Milkereit, P. & Tschochner, H. A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription. *The EMBO journal* **17,** 3692–3703; 10.1093/emboj/17.13.3692. (1998).

95. Peyroche, G. *et al.* The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3. *The EMBO journal* **19,** 5473–5482; 10.1093/emboj/19.20.5473 (2000).

96. Schnapp, A. & Grummt, I. Transcription complex formation at the mouse rDNA promoter involves the stepwise association of four transcription factors and RNA polymerase I. *Journal of Biological Chemistry* **266,** 24588–24595 (1991).

97. Schnapp, A., Schnapp, G., Erny, B. & Grummt, I. Function of the growth-regulated transcription initiation factor TIF-IA in initiation complex formation at the murine ribosomal gene promoter. *Molecular and cellular biology* **13,** 6723–6732; 10.1128/mcb.13.11.6723-6732.1993 (1993).

98. Stepanchick, A. *et al.* DNA binding by the ribosomal DNA transcription factor rrn3 is essential for ribosomal DNA transcription. *The Journal of biological chemistry* **288,** 9135–9144; 10.1074/jbc.M112.444265 (2013).

99. Gorski, J. J. *et al.* A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription. *The EMBO journal* **26,** 1560–1568; 10.1038/sj.emboj.7601601 (2007).

100. Heix, J., Zomerdijk, J. C., Ravanpay, A., Tjian, R. & Grummt, I. Cloning of murine RNA polymerase I-specific TAF factors: conserved interactions between the subunits of the species-specific transcription initiation factor TIF-IB/SL1. *Proceedings of the National Academy of Sciences of the United States of America* **94,** 1733–1738; 10.1073/pnas.94.5.1733 (1997).

101. Denissov, S. *et al.* Identification of novel functional TBP-binding sites and general factor repertoires. *The EMBO journal* **26,** 944–954; 10.1038/sj.emboj.7601550 (2007).

102. Panov, K. I. *et al.* RNA polymerase I-specific subunit CAST/hPAF49 has a role in the activation of transcription by upstream binding factor. *Molecular and cellular biology* **26,** 5436–5448; 10.1128/MCB.00230-06. (2006).

103. Miller, G. *et al.* hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters. *The EMBO journal* **20,** 1373–1382; 10.1093/emboj/20.6.1373 (2001).

104. Zomerdijk, J. C., Beckmann, H., Comai, L. & Tjian, R. Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits. *Science (New York, N.Y.)* **266,** 2015–2018; 10.1126/science.7801130 (1994).

105. Murano, K. *et al.* Reconstitution of human rRNA gene transcription in mouse cells by a complete SL1 complex. *Journal of cell science* **127,** 3309–3319; 10.1242/jcs.146787 (2014).

106. Bell, S. P., Jantzen, H. M. & Tjian, R. Assembly of alternative multiprotein complexes directs rRNA promoter selectivity. *Genes & development* **4,** 943–954; 10.1101/gad.4.6.943 (1990).

107. Friedrich, J. K., Panov, K. I., Cabart, P., Russell, J. & Zomerdijk, J. C. B. M. TBP-TAF complex SL1 directs RNA polymerase I pre-initiation complex formation and stabilizes upstream binding factor at the rDNA promoter. *Journal of Biological Chemistry* **280,** 29551–29558; 10.1074/jbc.M501595200 (2005).

108. Jantzen, H. M., Chow, A. M., King, D. S. & Tjian, R. Multiple domains of the RNA polymerase I activator hUBF interact with the TATA-binding protein complex hSL1 to mediate transcription. *Genes & development* **6,** 1950–1963; 10.1101/gad.6.10.1950 (1992).

109. McStay, B., Frazier, M. W. & Reeder, R. H. xUBF contains a novel dimerization domain essential for RNA polymerase I transcription. *Genes & development* **5,** 1957–1968; 10.1101/gad.5.11.1957 (1991).

110. O'Sullivan, A. C., Sullivan, G. J. & McStay, B. UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat. *Molecular and cellular biology* **22,** 657–668; 10.1128/MCB.22.2.657-668.2002 (2002).

111. Hamdane, N. *et al.* Conditional inactivation of Upstream Binding Factor reveals its epigenetic functions and the existence of a somatic nucleolar precursor body. *PLoS genetics* **10,** e1004505; 10.1371/journal.pgen.1004505 (2014).

112. Stefanovsky, V., Langlois, F., Gagnon-Kugler, T., Rothblum, L. I. & Moss, T. Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling. *Molecular cell* **21,** 629–639; 10.1016/j.molcel.2006.01.023 (2006).

113. Jantzen, H. M., Admon, A., Bell, S. P. & Tjian, R. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344,** 830–836; 10.1038/344830a0 (1990).

114. Neil, K. J. *et al.* Structure of recombinant rat UBF by electron image analysis and homology modelling. *Nucleic acids research* **24,** 1472–1480; 10.1093/nar/24.8.1472 (1996).

115. Putnam, C. D., Copenhaver, G. P., Denton, M. L. & Pikaard, C. S. The RNA polymerase I transactivator upstream binding factor requires its dimerization domain and high-mobility-group (HMG) box 1 to bend, wrap, and positively supercoil enhancer DNA. *Molecular and cellular biology* **14,** 6476–6488; 10.1128/mcb.14.10.6476-6488.1994 (1994).

116. Stefanovsky, V. Y., Pelletier, G., Bazett-Jones, D. P., Crane-Robinson, C. & Moss, T. DNA looping in the RNA polymerase I enhancesome is the result of non-cooperative in-phase bending by two UBF molecules. *Nucleic acids research* **29,** 3241–3247; 10.1093/nar/29.15.3241 (2001).

117. Kwon, H. & Green, M. R. The RNA polymerase I transcription factor, upstream binding factor, interacts directly with the TATA box-binding protein. *The Journal of biological chemistry* **269,** 30140–30146 (1994).

118. Tuan, J. C., Zhai, W. & Comai, L. Recruitment of TATA-binding protein-TAFI complex SL1 to the human ribosomal DNA promoter is mediated by the carboxy-terminal activation domain of upstream binding factor (UBF) and is regulated by UBF phosphorylation. *Molecular and cellular biology* **19,** 2872–2879; 10.1128/MCB.19.4.2872 (1999).

119. Voit, R. *et al.* The nucleolar transcription factor mUBF is phosphorylated by casein kinase II in the C-terminal hyperacidic tail which is essential for transactivation. *The EMBO journal* **11,** 2211–2218; 10.1002/j.1460-2075.1992.tb05280.x (1992).

120. Panov, K. I., Friedrich, J. K., Russell, J. & Zomerdijk, J. C. B. M. UBF activates RNA polymerase I transcription by stimulating promoter escape. *The EMBO journal* **25,** 3310–3322; 10.1038/sj.emboj.7601221 (2006).

121. O'Mahony, D. J. & Rothblum, L. I. Identification of two forms of the RNA polymerase I transcription factor UBF. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 3180–3184; 10.1073/pnas.88.8.3180 (1991).

122. Gerber, J. K. *et al.* Termination of mammalian rDNA replication: polar arrest of replication fork movement by transcription termination factor TTF-I. *Cell* **90,** 559–567; 10.1016/S0092-8674(00)80515-2 (1997).

123. Akamatsu, Y. & Kobayashi, T. The Human RNA Polymerase I Transcription Terminator Complex Acts as a Replication Fork Barrier That Coordinates the Progress of Replication with rRNA Transcription Activity. *Molecular and cellular biology* **35,** 1871–1881; 10.1128/MCB.01521-14 (2015).

124. Längst, G., Blank, T. A., Becker, P. B. & Grummt, I. RNA polymerase I transcription on nucleosomal templates: the transcription termination factor TTF-I induces chromatin remodeling and relieves transcriptional repression. *The EMBO journal* **16,** 760–768; 10.1093/emboj/16.4.760 (1997).

125. Längst, G., Becker, P. B. & Grummt, I. TTF-I determines the chromatin architecture of the active rDNA promoter. *The EMBO journal* **17,** 3135–3145; 10.1093/emboj/17.11.3135 (1998).

126. Strohner, R. *et al.* Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin. *Molecular and cellular biology* **24,** 1791–1798; 10.1128/MCB.24.4.1791-1798.2004 (2004).

127. Santoro, R., Li, J. & Grummt, I. The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat Genet* **32,** 393–396; 10.1038/ng1010 (2002).

128. Merkl, P. *et al.* Binding of the termination factor Nsi1 to its cognate DNA site is sufficient to terminate RNA polymerase I transcription in vitro and to induce termination in vivo. *Molecular and cellular biology* **34,** 3817–3827; 10.1128/MCB.00395-14 (2014).

129. Jansa, P., Mason, S. W., Hoffmann-Rohrer, U. & Grummt, I. Cloning and functional characterization of PTRF, a novel protein which induces dissociation of paused ternary transcription complexes. *The EMBO journal* **17,** 2855–2864; 10.1093/emboj/17.10.2855 (1998).

130. Zhang, Y., Sikes, M. L., Beyer, A. L. & Schneider, D. A. The Paf1 complex is required for efficient transcription elongation by RNA polymerase I. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 2153–2158; 10.1073/pnas.0812939106 (2009).

131. Birch, J. L. *et al.* FACT facilitates chromatin transcription by RNA polymerases I and III. *The EMBO journal* **28,** 854–865; 10.1038/emboj.2009.33 (2009).

132. Sadian, Y. *et al.* Structural insights into transcription initiation by yeast RNA polymerase I. *The EMBO journal* **36,** 2698–2709; 10.15252/embj.201796958 (2017).

133. Pilsl, M. & Engel, C. Structural basis of RNA polymerase I pre-initiation complex formation and promoter melting. *Nature communications* **11,** 1206; 10.1038/s41467-020-15052-y (2020).

134. Kostrewa, D., Kuhn, C.-D., Engel, C. & Cramer, P. An alternative RNA polymerase I structure reveals a dimer hinge. *Acta crystallographica. Section D, Biological crystallography* **71,** 1850–1855; 10.1107/S1399004715012651 (2015).

135. Bischler, N. *et al.* Localization of the yeast RNA polymerase I-specific subunits. *The EMBO journal* **21,** 4136–4144; 10.1093/emboj/cdf392 (2002).

136. Milkereit, P., Schultz, P. & Tschochner, H. Resolution of RNA polymerase I into dimers and monomers and their function in transcription. *Biological chemistry* **378,** 1433–1443; 10.1515/bchm.1997.378.12.1433 (1997).

137. Cramer, P. *et al.* Structure of eukaryotic RNA polymerases. *Annual review of biophysics* **37,** 337–352; 10.1146/annurev.biophys.37.032807.130008 (2008).

138. Lisica, A. *et al.* Mechanisms of backtrack recovery by RNA polymerases I and II. *Proceedings of the National Academy of Sciences of the United States of America* **113,** 2946–2951; 10.1073/pnas.1517011113 (2016).

139. Prescott, E. M. *et al.* Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 6068–6073; 10.1073/pnas.0401393101. (2004).

140. Lang, W. H., Morrow, B. E., Ju, Q., Warner, J. R. & Reeder, R. H. A model for transcription termination by RNA polymerase I. *Cell* **79,** 527–534; 10.1016/0092-8674(94)90261-5 (1994).

141. Reiter, A. *et al.* The Reb1-homologue Ydr026c/Nsi1 is required for efficient RNA polymerase I termination in yeast. *The EMBO journal* **31,** 3480–3493; 10.1038/emboj.2012.185 (2012).

142. Bartsch, I., Schoneberg, C. & Grummt, I. Purification and characterization of TTFI, a factor that mediates termination of mouse ribosomal DNA transcription. *Molecular and cellular biology* **8,** 3891–3897; 10.1128/mcb.8.9.3891-3897.1988 (1988).

143. Evers, R., Smid, A., Rudloff, U., Lottspeich, F. & Grummt, I. Different domains of the murine RNA polymerase I-specific termination factor mTTF-I serve distinct functions in transcription termination. *The EMBO journal* **14,** 1248–1256; 10.1002/j.1460-2075.1995.tb07108.x (1995).

144. Stefanovsky, V. & Moss, T. Regulation of rRNA synthesis in human and mouse cells is not determined by changes in active gene count. *Cell cycle (Georgetown, Tex.)* **5,** 735–739; 10.4161/cc.5.7.2633 (2006).

145. French, S. L., Osheim, Y. N., Cioci, F., Nomura, M. & Beyer, A. L. In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Molecular and cellular biology* **23,** 1558–1568; 10.1128/MCB.23.5.1558-1568.2003 (2003).

146. Bersaglieri, C. & Santoro, R. Genome Organization in and around the Nucleolus. *Cells* **8**; 10.3390/cells8060579 (2019).

147. Ide, S., Miyazaki, T., Maki, H. & Kobayashi, T. Abundance of ribosomal RNA gene copies maintains genome integrity. *Science (New York, N.Y.)* **327,** 693–696; 10.1126/science.1179044 (2010).

148. Hori, Y., Engel, C. & Kobayashi, T. Regulation of ribosomal RNA gene copy number, transcription and nucleolus organization in eukaryotes. *Nature reviews. Molecular cell biology*; 10.1038/s41580-022-00573-9 (2023).

149. Gadal, O. *et al.* A34.5, a nonessential component of yeast RNA polymerase I, cooperates with subunit A14 and DNA topoisomerase I to produce a functional rRNA synthesis machine. *Molecular and cellular biology* **17,** 1787–1795; 10.1128/MCB.17.4.1787 (1997).

150. Panova, T. B., Panov, K. I., Russell, J. & Zomerdijk, J. C. B. M. Casein kinase 2 associates with initiation-competent RNA polymerase I and has multiple roles in ribosomal DNA transcription. *Molecular and cellular biology* **26,** 5957–5968; 10.1128/MCB.00673-06 (2006).

151. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 10762–10767; 10.1073/pnas.0805139105 (2008).

152. Olsen, J. V. *et al.* Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648; 10.1016/j.cell.2006.09.026 (2006).

153. Olsen, J. V. *et al.* Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science signaling* **3,** ra3; 10.1126/scisignal.2000475 (2010).

154. Rigbolt, K. T. G. *et al.* System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. *Science signaling* **4,** rs3; 10.1126/scisignal.2001570 (2011).

155. Mayya, V. *et al.* Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. *Science signaling* **2,** ra46; 10.1126/scisignal.2000007 (2009).

156. Kamada, Y., Jung, U. S., Piotrowski, J. & Levin, D. E. The protein kinase C-activated MAP kinase pathway of Saccharomyces cerevisiae mediates a novel aspect of the heat shock response. *Genes & development* **9,** 1559–1571; 10.1101/gad.9.13.1559 (1995).

157. Voit, R., Kuhn, A., Sander, E. E. & Grummt, I. Activation of mammalian ribosomal gene transcription requires phosphorylation of the nucleolar transcription factor UBF. *Nucleic acids research* **23,** 2593–2599; 10.1093/nar/23.14.2593 (1995).

158. Zhao, J., Yuan, X., Frödin, M. & Grummt, I. ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth. *Molecular cell* **11,** 405–413; 10.1016/s1097-2765(03)00036-4 (2003).

159. Mayer, C., Zhao, J., Yuan, X. & Grummt, I. mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes & development* **18,** 423–434; 10.1101/gad.285504 (2004).

160. Mayer, C., Bierhoff, H. & Grummt, I. The nucleolus as a stress sensor: JNK2 inactivates the transcription factor TIF-IA and down-regulates rRNA synthesis. *Genes & development* **19,** 933–941; 10.1101/gad.333205 (2005).

161. Hoppe, S. *et al.* AMP-activated protein kinase adapts rRNA synthesis to cellular energy supply. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 17781–17786; 10.1073/pnas.0909873106 (2009).

162. Kim, J. K. *et al.* Sirtuin7 oncogenic potential in human hepatocellular carcinoma and its regulation by the tumor suppressors MiR-125a-5p and MiR-125b. *Hepatology (Baltimore, Md.)* **57,** 1055–1067; 10.1002/hep.26101 (2013).

163. Antony, C. *et al.* Control of ribosomal RNA synthesis by hematopoietic transcription factors. *Molecular cell* **82,** 3826-3839.e9; 10.1016/j.molcel.2022.08.027 (2022).

164. Claypool, J. A. *et al.* Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. *Molecular biology of the cell* **15,** 946–956; 10.1091/mbc.e03-08-0594. (2004).

165. Drygin, D., Rice, W. G. & Grummt, I. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annual review of pharmacology and toxicology* **50,** 131–156; 10.1146/annurev.pharmtox.010909.105844. (2010).

166. Grummt, I. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes & development* **17,** 1691–1702; 10.1101/gad.1098503R (2003).

167. Ferreira, R., Schneekloth, J. S., Panov, K. I., Hannan, K. M. & Hannan, R. D. Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age. *Cells* **9**; 10.3390/cells9020266 (2020).

168. Grandori, C. *et al.* c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nature cell biology* **7,** 311–318; 10.1038/ncb1224 (2005).

169. Arabi, A. *et al.* c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. *Nature cell biology* **7,** 303–310; 10.1038/ncb1225 (2005).

170. Bierhoff, H., Dundr, M., Michels, A. A. & Grummt, I. Phosphorylation by casein kinase 2 facilitates rRNA gene transcription by promoting dissociation of TIF-IA from elongating RNA polymerase I. *Molecular and cellular biology* **28,** 4988–4998; 10.1128/MCB.00492-08 (2008).

171. Lin, C.-Y., Navarro, S., Reddy, S. & Comai, L. CK2-mediated stimulation of Pol I transcription by stabilization of UBF-SL1 interaction. *Nucleic acids research* **34,** 4752–4766; 10.1093/nar/gkl581 (2006).

172. Duployez, N. *et al.* UBTF tandem duplications define a distinct subtype of adult de novo acute myeloid leukemia. *Leukemia* **37,** 1245–1253; 10.1038/s41375-023-01906-z (2023).

173. Umeda, M. *et al.* Integrated Genomic Analysis Identifies UBTF Tandem Duplications as a Recurrent Lesion in Pediatric Acute Myeloid Leukemia. *Blood cancer discovery* **3,** 194–207; 10.1158/2643-3230.BCD-21-0160 (2022).

174. Andrews, W. J., Ray, S., Panova, T., Engel, C. & Panov, K. I. DNA Intercalators Inhibit Eukaryotic Ribosomal RNA Synthesis by Impairing the Initiation of Transcription. *Genes* **12**; 10.3390/genes12091412 (2021).

175. Mars, J.-C. *et al.* The chemotherapeutic agent CX-5461 irreversibly blocks RNA polymerase I initiation and promoter release to cause nucleolar disruption, DNA damage and cell inviability. *NAR cancer* **2,** zcaa032; 10.1093/narcan/zcaa032 (2020).

176. Daiß, J. L., Griesenbeck, J., Tschochner, H. & Engel, C. Synthesis of the ribosomal RNA precursor in human cells: mechanisms, factors and regulation. *Biological chemistry*; 10.1515/hsz-2023-0214 (2023).

177. McNamar, R. *et al.* PAF49: An RNA Polymerase I subunit essential for rDNA transcription and stabilization of PAF53. *The Journal of biological chemistry* **299,** 104951; 10.1016/j.jbc.2023.104951 (2023).

178. Knutson, B. A., McNamar, R. & Rothblum, L. I. Dynamics of the RNA polymerase I TFIIF/TFIIE-like subcomplex: a mini-review. *Biochemical Society transactions* **48,** 1917–1927; 10.1042/BST20190848 (2020).

179. Ramsay, E. P. *et al.* Structure of human RNA polymerase III. *Nature communications* **11,** 6409; 10.1038/s41467-020-20262-5 (2020).

180. Gunkel, P., Iino, H., Krull, S. & Cordes, V. C. ZC3HC1 Is a Novel Inherent Component of the Nuclear Basket, Resident in a State of Reciprocal Dependence with TPR. *Cells* **10**; 10.3390/cells10081937 (2021).

181. Scheer, U. & Rose, K. M. Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. *Proceedings of the National Academy of Sciences of the United States of America* **81,** 1431–1435; 10.1073/pnas.81.5.1431 (1984).

182. Dimario, P. J. Cell and molecular biology of nucleolar assembly and disassembly. *International review of cytology* **239,** 99–178; 10.1016/S0074-7696(04)39003-0 (2004).

183. Andrews, J. O. *et al.* qSR: a quantitative super-resolution analysis tool reveals the cell-cycle dependent organization of RNA Polymerase I in live human cells. *Scientific reports* **8,** 7424; 10.1038/s41598-018-25454-0 (2018).

184. Engel, C. Purification of Crystallization-Grade RNA Polymerase I from S. cerevisiae. *Methods in molecular biology (Clifton, N.J.)* **1455,** 85–97; 10.1007/978-1-4939-3792-9_7. (2016).

185. Smid, A., Riva, M., Bouet, F., Sentenac, A. & Carles, C. The association of three subunits with yeast RNA polymerase is stabilized by A14. *The Journal of biological chemistry* **270,** 13534–13540; 10.1074/jbc.270.22.13534 (1995).

186. Jacobs, R. Q., Ingram, Z. M., Lucius, A. L. & Schneider, D. A. Defining the divergent enzymatic properties of RNA polymerases I and II. *The Journal of biological chemistry* **296,** 100051; 10.1074/jbc.RA120.015904 (2021).

187. Pilsl, M. & Engel, C. Structural Studies of Eukaryotic RNA Polymerase I Using Cryo-Electron Microscopy. *Methods in molecular biology (Clifton, N.J.)* **2533,** 71–80; 10.1007/978-1-0716-2501-9_5 (2022).

188. Pilsl, M. *et al.* Preparation of RNA Polymerase Complexes for Their Analysis by Single-Particle Cryo-Electron Microscopy. *Methods in molecular biology (Clifton, N.J.)* **2533,** 81–96; 10.1007/978-1-0716-2501-9_6. (2022).

189. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nature methods* **16,** 1146–1152; 10.1038/s41592-019-0580-y (2019).

190. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**; 10.7554/eLife.42166 (2018).

191. Girbig, M. *et al.* Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. *Nature structural & molecular biology* **28,** 210–219; 10.1038/s41594-020-00555-5 (2021).

192. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Current protocols in protein science* **86,** 2.9.1-2.9.37; 10.1002/cpps.20 (2016).

193. 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; 10.1107/S2059798319011471 (2019).

194. Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* **596,** 590–596; 10.1038/s41586-021-03828-1 (2021).

195. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of molecular biology* **430,** 2237–2243; 10.1016/j.jmb.2017.12.007 (2018).

196. Weaver, K. N. *et al.* Acrofacial Dysostosis, Cincinnati Type, a Mandibulofacial Dysostosis Syndrome with Limb Anomalies, Is Caused by POLR1A Dysfunction. *American journal of human genetics* **96,** 765–774; 10.1016/j.ajhg.2015.03.011 (2015).

197. Ide, S., Imai, R., Ochi, H. & Maeshima, K. Transcriptional suppression of ribosomal DNA with phase separation. *Science advances* **6**; 10.1126/sciadv.abb5953 (2020).

198. Schaefer, E. *et al.* Autosomal recessive POLR1D mutation with decrease of TCOF1 mRNA is responsible for Treacher Collins syndrome. *Genetics in medicine : official journal of the American College of Medical Genetics* **16,** 720–724; 10.1038/gim.2014.12 (2014).

199. 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; 10.1038/ng.724 (2011).

200. Thiffault, I. *et al.* Recessive mutations in POLR1C cause a leukodystrophy by impairing biogenesis of RNA polymerase III. *Nature communications* **6,** 7623; 10.1038/ncomms8623 (2015).

201. Sanchez, E. *et al.* POLR1B and neural crest cell anomalies in Treacher Collins syndrome type 4. *Genetics in medicine : official journal of the American College of Medical Genetics* **22,** 547–556; 10.1038/s41436-019-0669-9 (2020).

202. Gauquelin, L. *et al.* Clinical spectrum of POLR3-related leukodystrophy caused by biallelic POLR1C pathogenic variants. *Neurology. Genetics* **5,** e369; 10.1212/NXG.0000000000000369 (2019).

203. Kara, B. *et al.* Severe neurodegenerative disease in brothers with homozygous mutation in POLR1A. *European journal of human genetics : EJHG* **25,** 315–323; 10.1038/ejhg.2016.183 (2017).

204. García-López, M. C. *et al.* The conserved foot domain of RNA pol II associates with proteins involved in transcriptional initiation and/or early elongation. *Genetics* **189,** 1235–1248; 10.1534/genetics.111.133215 (2011).

205. Plaschka, C. *et al.* Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature* **518,** 376–380; 10.1038/nature14229 (2015).

206. Stros, M., Launholt, D. & Grasser, K. D. The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cellular and molecular life sciences : CMLS* **64,** 2590–2606; 10.1007/s00018-007-7162-3 (2007).

207. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein science : a publication of the Protein Society* **27,** 14–25; 10.1002/pro.3235 (2018).

208. Kim, M.-S. *et al.* Cracking the DNA Code for V(D)J Recombination. *Molecular cell* **70,** 358-370.e4; 10.1016/j.molcel.2018.03.008 (2018).

209. Cary, P. D., Turner, C. H., Mayes, E. & Crane-Robinson, C. Conformation and domain structure of the non-histone chromosomal proteins, HMG 1 and 2. Isolation of two folded fragments from HMG 1 and 2. *European journal of biochemistry* **131,** 367–374; 10.1111/j.1432-1033.1983.tb07272.x (1983).

210. Ray, S. *et al.* Topoisomerase IIα promotes activation of RNA polymerase I transcription by facilitating pre-initiation complex formation. *Nature communications* **4,** 1598; 10.1038/ncomms2599 (2013).

211. Stros, M., Bacíková, A., Polanská, E., Stokrová, J. & Strauss, F. HMGB1 interacts with human topoisomerase IIalpha and stimulates its catalytic activity. *Nucleic acids research* **35,** 5001–5013; 10.1093/nar/gkm525 (2007).

212. Vanden Broeck, A. *et al.* Structural basis for allosteric regulation of Human Topoisomerase IIα. *Nature communications* **12,** 2962; 10.1038/s41467-021-23136-6 (2021).

213. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society* **125,** 1731–1737; 10.1021/ja026939x (2003).

214. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31,** 455–461; 10.1002/jcc.21334 (2010).

215. Pierce, B. G. *et al.* ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics (Oxford, England)* **30,** 1771–1773; 10.1093/bioinformatics/btu097 (2014).

216. Baspinar, A., Cukuroglu, E., Nussinov, R., Keskin, O. & Gursoy, A. PRISM: a web server and repository for prediction of protein-protein interactions and modeling their 3D complexes. *Nucleic acids research* **42,** W285-9; 10.1093/nar/gku397 (2014).

217. Tuncbag, N., Gursoy, A., Nussinov, R. & Keskin, O. Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. *Nature protocols* **6,** 1341–1354; 10.1038/nprot.2011.367 (2011).

218. Scheurer, M. *et al.* PyContact: Rapid, Customizable, and Visual Analysis of Noncovalent Interactions in MD Simulations. *Biophysical journal* **114,** 577–583; 10.1016/j.bpj.2017.12.003 (2018).

219. Canela, A. *et al.* Genome Organization Drives Chromosome Fragility. *Cell* **170,** 507-521.e18; 10.1016/j.cell.2017.06.034 (2017).

220. Naidu, S., Friedrich, J. K., Russell, J. & Zomerdijk, J. C. B. M. TAF1B is a TFIIB-like component of the basal transcription machinery for RNA polymerase I. *Science (New York, N.Y.)* **333,** 1640–1642; 10.1126/science.1207656 (2011).

221. Mars, J.-C., Sabourin-Felix, M., Tremblay, M. G. & Moss, T. A Deconvolution Protocol for ChIP-Seq Reveals Analogous Enhancer Structures on the Mouse and Human Ribosomal RNA Genes. *G3 (Bethesda, Md.)* **8,** 303–314; 10.1534/g3.117.300225 (2018).

222. Misiaszek, A. D. *et al.* Cryo-EM structures of human RNA polymerase I. *Nature structural & molecular biology* **28,** 997–1008; 10.1038/s41594-021-00693-4 (2021).

223. Zhao, D. *et al.* Structure of the human RNA polymerase I elongation complex. *Cell discovery* **7,** 97; 10.1038/s41421-021-00335-5 (2021).

224. Daiß, J. L. *et al.* The human RNA polymerase I structure reveals an HMG-like docking domain specific to metazoans. *Life science alliance* **5**; 10.26508/lsa.202201568 (2022).

225. Maiser, A. *et al.* Super-resolution in situ analysis of active ribosomal DNA chromatin organization in the nucleolus. *Scientific reports* **10,** 7462; 10.1038/s41598-020-64589-x (2020).

226. Rossi, M. J. *et al.* A high-resolution protein architecture of the budding yeast genome. *Nature,* 1–6; 10.1038/s41586-021-03314-8 (2021).

227. Hanada, K. *et al.* RNA polymerase I associated factor 53 binds to the nucleolar transcription factor UBF and functions in specific rDNA transcription. *The EMBO journal* **15,** 2217–2226 (1996).

228. Dundr, M. *et al.* A kinetic framework for a mammalian RNA polymerase in vivo. *Science (New York, N.Y.)* **298,** 1623–1626; 10.1126/science.1076164 (2002).

229. Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M. & Cramer, P. Structure of transcribing mammalian RNA polymerase II. *Nature* **529,** 551–554; 10.1038/nature16482 (2016).

230. Lee, J. H. & Berger, J. M. Cell Cycle-Dependent Control and Roles of DNA Topoisomerase II. *Genes* **10**; 10.3390/genes10110859 (2019).

231. Morotomi-Yano, K. & Yano, K.-I. Nucleolar translocation of human DNA topoisomerase II by ATP depletion and its disruption by the RNA polymerase I inhibitor BMH-21. *Scientific reports* **11,** 21533; 10.1038/s41598-021-00958-4 (2021).

232. Schilbach, S., Aibara, S., Dienemann, C., Grabbe, F. & Cramer, P. Structure of RNA polymerase II pre-initiation complex at 2.9 Å defines initial DNA opening. *Cell* **184,** 4064-4072.e28; 10.1016/j.cell.2021.05.012 (2021).

233. Aibara, S., Schilbach, S. & Cramer, P. Structures of mammalian RNA polymerase II pre-initiation complexes. *Nature* **594,** 124–128; 10.1038/s41586-021-03554-8 (2021).

234. Abdella, R. *et al.* Structure of the human Mediator-bound transcription preinitiation complex. *Science (New York, N.Y.)* **372,** 52–56; 10.1126/science.abg3074 (2021).

235. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America* **84,** 7024–7027; 10.1073/pnas.84.20.7024 (1987).

236. French, S. L. *et al.* Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes. *Molecular and cellular biology* **31,** 482–494; 10.1128/MCB.00589-10 (2011).

237. Onoda, A. *et al.* Nuclear dynamics of topoisomerase IIβ reflects its catalytic activity that is regulated by binding of RNA to the C-terminal domain. *Nucleic acids research* **42,** 9005–9020; 10.1093/nar/gku640 (2014).

238. Christensen, M. O. *et al.* Dynamics of human DNA topoisomerases IIalpha and IIbeta in living cells. *The Journal of cell biology* **157,** 31–44; 10.1083/jcb.200112023 (2002).

239. Colis, L. *et al.* DNA intercalator BMH-21 inhibits RNA polymerase I independent of DNA damage response. *Oncotarget* **5,** 4361–4369; 10.18632/oncotarget.2020 (2014).

240. Musso, L. *et al.* c-MYC G-quadruplex binding by the RNA polymerase I inhibitor BMH-21 and analogues revealed by a combined NMR and biochemical Approach. *Biochimica et biophysica acta. General subjects* **1862,** 615–629; 10.1016/j.bbagen.2017.12.002 (2018).

241. Jacobs, R. Q. *et al.* RNA Polymerase I Is Uniquely Vulnerable to the Small-Molecule Inhibitor BMH-21. *Cancers* **14**; 10.3390/cancers14225544 (2022).

242. Wei, T. *et al.* Small-Molecule Targeting of RNA Polymerase I Activates a Conserved Transcription Elongation Checkpoint. *Cell reports* **23,** 404–414; 10.1016/j.celrep.2018.03.066 (2018).

243. Jacobs, R. Q., Huffines, A. K., Laiho, M. & Schneider, D. A. The small-molecule BMH-21 directly inhibits transcription elongation and DNA occupancy of RNA polymerase I in vivo and in vitro. *The Journal of biological chemistry* **298,** 101450; 10.1016/j.jbc.2021.101450 (2022).

244. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nature biotechnology* **22,** 1583–1587; 10.1038/nbt1036 (2004).

245. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8,** 2281–2308; 10.1038/nprot.2013.143 (2013).

246. Eulalio, A., Huntzinger, E. & Izaurralde, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nature structural & molecular biology* **15,** 346–353; 10.1038/nsmb.1405 (2008).

247. Davis, M. W. & Jorgensen, E. M. ApE, A Plasmid Editor: A Freely Available DNA Manipulation and Visualization Program. *Frontiers in bioinformatics* **2,** 818619; 10.3389/fbinf.2022.818619 (2022).

248. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of structural biology* **152,** 36–51; 10.1016/j.jsb.2005.07.007 (2005).

249. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* **180,** 519–530; 10.1016/j.jsb.2012.09.006 (2012).

250. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25,** 1605–1612; 10.1002/jcc.20084. (2004).

251. Bond, C. S. & Schüttelkopf, A. W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta crystallographica. Section D, Biological crystallography* **65,** 510–512; 10.1107/S0907444909007835 (2009).

252. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66,** 213–221; 10.1107/S0907444909052925 (2010).

253. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66,** 486–501; 10.1107/S0907444910007493 (2010).

254. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook,* edited by J. M. Walker (Humana Press, Totowa, NJ, 2005), pp. 571–607.

255. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nature methods* **3,** 1021–1032; 10.1038/nmeth983 (2006).

256. Ma, H. T. & Poon, R. Y. C. Synchronization of HeLa Cells. *Methods in molecular biology (Clifton, N.J.)* **1524,** 189–201; 10.1007/978-1-4939-6603-5_12 (2017).

257. Bedez, C. *et al.* Post-translational modifications in DNA topoisomerase 2α highlight the role of a eukaryote-specific residue in the ATPase domain. *Scientific reports* **8,** 9272; 10.1038/s41598-018-27606-8 (2018).

258. Chen, S.-F. *et al.* Structural insights into the gating of DNA passage by the topoisomerase II DNA-gate. *Nature communications* **9,** 3085; 10.1038/s41467-018-05406-y (2018).

259. Thomas G. Martin, Andreas Boland, Anthony W. P. Fitzpatrick & Sjors H. W. Scheres. *Graphene Oxide Grid Preparation* (figshare, 2016).

260. Palovcak, E. *et al.* A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *Journal of structural biology* **204,** 80–84; 10.1016/j.jsb.2018.07.007 (2018).

261. Patel, A., Toso, D., Litvak, A. & Nogales, E. *Efficient graphene oxide coating improves cryo-EM sample preparation and data collection from tilted grids* (2021).

262. Schorb, M., Haberbosch, I., Hagen, W. J. H., Schwab, Y. & Mastronarde, D. N. Software tools for automated transmission electron microscopy. *Nature methods* **16,** 471–477; 10.1038/s41592-019-0396-9 (2019).

263. Fislage, M., Shkumatov, A. V., Stroobants, A. & Efremov, R. G. Assessing the JEOL CRYO ARM 300 for high-throughput automated single-particle cryo-EM in a multiuser environment. *IUCrJ* **7,** 707–718; 10.1107/S2052252520006065 (2020).

264. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60,** 2126–2132; 10.1107/S0907444904019158 (2004).

265. Blum, M. *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic acids research* **49,** D344-D354; 10.1093/nar/gkaa977 (2021).

266. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30,** 772–780; 10.1093/molbev/mst010 (2013).

267. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution* **17,** 540–552; 10.1093/oxfordjournals.molbev.a026334 (2000).

268. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics (Oxford, England)* **22,** 2688–2690; 10.1093/bioinformatics/btl446 (2006).

269. Burki, F., Roger, A. J., Brown, M. W. & Simpson, A. G. B. The New Tree of Eukaryotes. *Trends in ecology & evolution* **35,** 43–55; 10.1016/j.tree.2019.08.008 (2020).

270. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)* **25,** 1189–1191; 10.1093/bioinformatics/btp033 (2009).

271. Gabler, F. *et al.* Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Current protocols in bioinformatics* **72,** e108; 10.1002/cpbi.108 (2020).

272. Mitternacht, S. FreeSASA: An open source C library for solvent accessible surface area calculations. *F1000Research* **5,** 189; 10.12688/f1000research.7931.1 (2016).

273. Geiduschek, E. P. & Kassavetis, G. A. The RNA polymerase III transcription apparatus. *Journal of molecular biology* **310,** 1–26; 10.1006/jmbi.2001.4732. (2001).

274. Filer, D. *et al.* RNA polymerase III limits longevity downstream of TORC1. *Nature* **552,** 263–267; 10.1038/nature25007 (2017).

275. White, R. J. Transcription by RNA polymerase III: more complex than we thought. *Nature reviews. Genetics* **12,** 459–463; 10.1038/nrg3001 (2011).

276. White, R. J. RNA polymerases I and III, non-coding RNAs and cancer. *Trends in genetics : TIG* **24,** 622–629; 10.1016/j.tig.2008.10.003 (2008).

277. White, R. J. RNA polymerase III transcription and cancer. *Oncogene* **23,** 3208–3216; 10.1038/sj.onc.1207547. (2004).

278. Paolacci, S. *et al.* Specific combinations of biallelic POLR3A variants cause Wiedemann-Rautenstrauch syndrome. *Journal of medical genetics* **55,** 837–846; 10.1136/jmedgenet-2018-105528 (2018).

279. Wambach, J. A. *et al.* Bi-allelic POLR3A Loss-of-Function Variants Cause Autosomal-Recessive Wiedemann-Rautenstrauch Syndrome. *American journal of human genetics* **103,** 968–975; 10.1016/j.ajhg.2018.10.010 (2018).

280. 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; 10.1016/j.ajhg.2011.07.014 (2011).

281. 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; 10.1016/j.ajhg.2011.10.006 (2011).

282. Saitsu, H. *et al.* Mutations in POLR3A and POLR3B encoding RNA Polymerase III subunits cause an autosomal-recessive hypomyelinating leukoencephalopathy. *American journal of human genetics* **89,** 644–651; 10.1016/j.ajhg.2011.10.003 (2011).

283. Ogunjimi, B. *et al.* Inborn errors in RNA polymerase III underlie severe varicella zoster virus infections. *The Journal of clinical investigation* **127,** 3543–3556; 10.1172/JCI92280 (2017).

284. Carter-Timofte, M. E., Paludan, S. R. & Mogensen, T. H. RNA Polymerase III as a Gatekeeper to Prevent Severe VZV Infections. *Trends in molecular medicine* **24,** 904–915; 10.1016/j.molmed.2018.07.009 (2018).

285. 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; 10.1038/s41435-018-0027-y (2019).

286. Hoffmann, N. A. *et al.* Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528,** 231–236; 10.1038/nature16143 (2015).

287. 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; 10.1016/j.molcel.2006.05.013. (2006).

288. Armache, K.-J., Mitterweger, S., Meinhart, A. & Cramer, P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *The Journal of biological chemistry* **280,** 7131–7134; 10.1074/jbc.M413038200 (2005).

289. Werner, M., Thuriaux, P. & Soutourina, J. Structure-function analysis of RNA polymerases I and III. *Current opinion in structural biology* **19,** 740–745; 10.1016/j.sbi.2009.10.005. (2009).

290. Whitehall, S. K., Bardeleben, C. & Kassavetis, G. A. Hydrolytic cleavage of nascent RNA in RNA polymerase III ternary transcription complexes. *The Journal of biological chemistry* **269,** 2299–2306 (1994).

291. Alic, N. *et al.* Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 10400–10405; 10.1073/pnas.0704116104 (2007).

292. Mishra, S., Hasan, S. H., Sakhawala, R. M., Chaudhry, S. & Maraia, R. J. Mechanism of RNA polymerase III termination-associated reinitiation-recycling conferred by the essential function of the N terminal-and-linker domain of the C11 subunit. *Nature communications* **12,** 5900; 10.1038/s41467-021-26080-7. (2021).

293. Chédin, 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; 10.1101/gad.12.24.3857 (1998).

294. 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; 10.1016/j.molcel.2007.02.016. (2007).

295. 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. *The Journal of biological chemistry* **285,** 2695–2706; 10.1074/jbc.M109.074013 (2010).

296. Arimbasseri, A. G., Rijal, K. & Maraia, R. J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5,** e27639; 10.4161/trns.27369 (2014).

297. 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; 10.1016/j.molcel.2015.04.002 (2015).

298. Arimbasseri, A. G. & Maraia, R. J. RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends in biochemical sciences* **41,** 546–559; 10.1016/j.tibs.2016.03.003 (2016).

299. Kessler, A. C. & Maraia, R. J. The nuclear and cytoplasmic activities of RNA polymerase III, and an evolving transcriptome for surveillance. *Nucleic acids research* **49,** 12017–12034; 10.1093/nar/gkab1145 (2021).

300. 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; 10.1016/s0092-8674(80)80048-1 (1980).

301. Wang, Q., Daiß, J. L., Xu, Y. & Engel, C. Snapshots of RNA polymerase III in action - A mini review. *Gene* **821,** 146282; 10.1016/j.gene.2022.146282 (2022).

302. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes & development* **16,** 2593–2620; 10.1101/gad.1018902. (2002).

303. Orioli, A., Pascali, C., Pagano, A., Teichmann, M. & Dieci, G. RNA polymerase III transcription control elements: themes and variations. *Gene* **493,** 185–194; 10.1016/j.gene.2011.06.015 (2012).

304. Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553,** 301–306; 10.1038/nature25441 (2018).

305. Liao, X. B., Clemens, K. R., Tennant, L., Wright, P. E. & Gottesfeld, J. M. Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the Xenopus 5 S RNA gene. *Journal of molecular biology* **223,** 857–871; 10.1016/0022-2836(92)90248-i. (1992).

306. Foster, M. P. *et al.* Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nature structural biology* **4,** 605–608; 10.1038/nsb0897-605. (1997).

307. Lu, D., Searles, M. A. & Klug, A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* **426,** 96–100; 10.1038/nature02088. (2003).

308. Lu, D. & Klug, A. Invariance of the zinc finger module: a comparison of the free structure with those in nucleic-acid complexes. *Proteins* **67,** 508–512; 10.1002/prot.21289 (2007).

309. Gouge, J. *et al.* Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. *Nature communications* **8,** 130; 10.1038/s41467-017-00126-1. (2017).

310. Gouge, J. *et al.* Redox Signaling by the RNA Polymerase III TFIIB-Related Factor Brf2. *Cell* **163,** 1375–1387; 10.1016/j.cell.2015.11.005 (2015).

311. Ramsay, E. P. & Vannini, A. Structural rearrangements of the RNA polymerase III machinery during tRNA transcription initiation. *Biochimica et biophysica acta. Gene regulatory mechanisms* **1861,** 285–294; 10.1016/j.bbagrm.2017.11.005 (2018).

312. Marzouki, N., Camier, S., Ruet, A., Moenne, A. & Sentenac, A. Selective proteolysis defines two DNA binding domains in yeast transcription factor tau. *Nature* **323,** 176–178; 10.1038/323176a0. (1986).

313. Lata, E. *et al.* RNA Polymerase III Subunit Mutations in Genetic Diseases. *Frontiers in molecular biosciences* **8,** 696438; 10.3389/fmolb.2021.696438 (2021).

314. Vorländer, M. K. *et al.* Structure of the TFIIIC subcomplex τA provides insights into RNA polymerase III pre-initiation complex formation. *Nature communications* **11,** 4905; 10.1038/s41467-020-18707-y. (2020).

315. Vannini, A. *et al.* Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell* **143,** 59–70; 10.1016/j.cell.2010.09.002. (2010).

316. Boguta, M. Maf1, a general negative regulator of RNA polymerase III in yeast. *Biochimica et biophysica acta* **1829,** 376–384; 10.1016/j.bbagrm.2012.11.004 (2013).

317. Boguta, M., Czerska, K. & Zoładek, T. Mutation in a new gene MAF1 affects tRNA suppressor efficiency in Saccharomyces cerevisiae. *Gene* **185,** 291–296; 10.1016/s0378-1119(96)00669-5 (1997).

318. Turowski, T. W. *et al.* Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome research* **26,** 933–944; 10.1101/gr.205492.116 (2016).

319. Vorländer, M. K. *et al.* Structural basis for RNA polymerase III transcription repression by Maf1. *Nature structural & molecular biology* **27,** 229–232; 10.1038/s41594-020-0383-y (2020).

320. 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 of the United States of America* **103,** 15044–15049; 10.1073/pnas.0607129103. (2006).

321. Lee, J., Moir, R. D. & Willis, I. M. Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. *The Journal of biological chemistry* **284,** 12604–12608; 10.1074/jbc.C900020200 (2009).

322. Graczyk, D. *et al.* Casein kinase II-mediated phosphorylation of general repressor Maf1 triggers RNA polymerase III activation. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 4926–4931; 10.1073/pnas.1010010108 (2011).

323. 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; 10.1016/j.molcel.2006.04.008 (2006).

324. Jarrous, N. & Rouvinski, A. RNA polymerase III and antiviral innate immune response. *Transcription* **12,** 1–11; 10.1080/21541264.2021.1890915 (2021).

325. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* **14,** 331–332; 10.1038/nmeth.4193 (2017).

326. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology* **192,** 216–221; 10.1016/j.jsb.2015.08.008 (2015).

327. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *eLife* **7**; 10.7554/eLife.35383 (2018).

328. 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; 10.1101/gad.11.10.1315 (1997).

329. Vos, S. M. *et al.* Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560,** 607–612; 10.1038/s41586-018-0440-4 (2018).

330. Vos, S. M., Farnung, L., Linden, A., Urlaub, H. & Cramer, P. Structure of complete Pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. *Nature structural & molecular biology* **27,** 668–677; 10.1038/s41594-020-0437-1 (2020).

331. Daoud, H. *et al.* Mutations in POLR3A and POLR3B are a major cause of hypomyelinating leukodystrophies with or without dental abnormalities and/or hypogonadotropic hypogonadism. *Journal of medical genetics* **50,** 194–197; 10.1136/jmedgenet-2012-101357 (2013).

332. Terao, Y. *et al.* Diffuse central hypomyelination presenting as 4H syndrome caused by compound heterozygous mutations in POLR3A encoding the catalytic subunit of polymerase III. *Journal of the neurological sciences* **320,** 102–105; 10.1016/j.jns.2012.07.005 (2012).

333. Shimojima, K. *et al.* Novel compound heterozygous mutations of POLR3A revealed by whole-exome sequencing in a patient with hypomyelination. *Brain & development* **36,** 315–321; 10.1016/j.braindev.2013.04.011 (2014).

334. Potic, A., Brais, B., Choquet, K., Schiffmann, R. & Bernard, G. 4H syndrome with late-onset growth hormone deficiency caused by POLR3A mutations. *Archives of neurology* **69,** 920–923; 10.1001/archneurol.2011.1963 (2012).

335. Carter-Timofte, M. E. *et al.* Varicella-zoster virus CNS vasculitis and RNA polymerase III gene mutation in identical twins. *Neurology(R) neuroimmunology & neuroinflammation* **5,** e500; 10.1212/NXI.0000000000000500 (2018).

336. Walker-Kopp, N. *et al.* Treacher Collins syndrome mutations in Saccharomyces cerevisiae destabilize RNA polymerase I and III complex integrity. *Human molecular genetics* **26,** 4290–4300; 10.1093/hmg/ddx317. (2017).

337. Wang, Q. *et al.* Structural insights into transcriptional regulation of human RNA polymerase III. *Nature structural & molecular biology* **28,** 220–227; 10.1038/s41594-021-00557-x (2021).

338. Li, L. *et al.* Structure of human RNA polymerase III elongation complex. *Cell research* **31,** 791–800; 10.1038/s41422-021-00472-2 (2021).

339. Hou, H. *et al.* Structural insights into RNA polymerase III-mediated transcription termination through trapping poly-deoxythymidine. *Nature communications* **12,** 6135; 10.1038/s41467-021-26402-9. (2021).

340. Cheng, R. & van Bortle, K. RNA polymerase III transcription and cancer: A tale of two RPC7 subunits. *Frontiers in molecular biosciences* **9,** 1073795; 10.3389/fmolb.2022.1073795 (2022).

341. 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 of the United States of America* **107,** 4176–4181; 10.1073/pnas.0914980107. (2010).

342. Smith, J. C. & Sheltzer, J. M. Genome-wide identification and analysis of prognostic features in human cancers. *Cell reports* **38,** 110569; 10.1016/j.celrep.2022.110569 (2022).

343. 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; 10.1002/j.1460-2075.1995.tb07009.x. (1995).

344. Xu, Y., Yang, W., Wu, J. & Shi, Y. Solution structure of the first HMG box domain in human upstream binding factor. *Biochemistry* **41,** 5415–5420; 10.1021/bi015977a (2002).

345. Yang, W., Xu, Y., Wu, J., Zeng, W. & Shi, Y. Solution structure and DNA binding property of the fifth HMG box domain in comparison with the first HMG box domain in human upstream binding factor. *Biochemistry* **42,** 1930–1938; 10.1021/bi026372x (2003).

346. Bazett-Jones, D. P., Leblanc, B., Herfort, M. & Moss, T. Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF. *Science (New York, N.Y.)* **264,** 1134–1137; 10.1126/science.8178172 (1994).

347. Nikolov, D. B. *et al.* Crystal structure of a human TATA box-binding protein/TATA element complex. *Proceedings of the National Academy of Sciences of the United States of America* **93,** 4862–4867; 10.1073/pnas.93.10.4862 (1996).

348. Juo, Z. S. *et al.* How proteins recognize the TATA box. *Journal of molecular biology* **261,** 239–254; 10.1006/jmbi.1996.0456 (1996).