UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS
DEPARTAMENTO DE BIOLOGÍA MOLECULAR

CRYO-EM STRUCTURES OF FREE MONOMERIC AND RRN3-BOUND RNA POLYMERASE I UNVEIL THE STRUCTURAL CHANGES IN THE TRANSITION FROM INACTIVE DIMERS TO THE ACTIVATED STATE

TESIS DOCTORAL

AUTOR:

Jaime Manuel Alegrio Louro

Licenciado en Farmacia

DIRECTOR:

Carlos Fernández-Tornero

Centro de Investigaciones Biológicas (CIB) - CSIC

Madrid, 2017

AGRADECIMIENTOS

A mi director de tesis, el Dr. Carlos Fernández-Tornero, por la oportunidad de realizar este trabajo, por la supervisión del mismo y por todo lo que me enseñó. Sin su contribución este trabajo no hubiera sido posible.

A la Dra. Eva Torreira por trabajar conmigo en el mismo proyecto y por las sugerencias meditadas, oportunas y con calidad científica superior. Eva participó en objetivos centrales de este trabajo y sin ella esta tesis hubiera tenido menos calidad.

Al Ministerio de Economía, Industria y Competitividad por la beca FPI BES-2011-044359.

A todos los compañeros de laboratorio que he tenido en estos cinco años y medio, de cada uno llevo algo para la vida. Gracias Rocío, Nuria, Marta, Eva, Nicholas, Marcel, Miguel, Alejandro y Margarita. Gracias a los compañeros del laboratorio vecino, Elena, Mercedes, Federico, Antonio, Irene, Javier y Andrea.

Al Servicio de Proteómica del CNB-CSIC, miembro de la plataforma ProteoRed-ISCIII, por el análisis proteómico a partir de geles nativos.

Al Servicio de Proteómica y Genómica del CIB-CSIC, miembro de la plataforma ProteoRed-ISCIII, por el análisis de masa molecular (MS-MALDI TOF) y la identificación de proteínas mediante MALDI-TOF-TOF.

Al Servicio de Química de Proteínas del CIB-CSIC por la secuenciación de proteínas.

Al Servicio de Ultracentrifugación Analítica y Dispersión de Luz del CIB-CSIC por los ensayos de velocidad de sedimentación.

Al Servicio de Fermentación del CBM-CSIC por las fermentaciones de levadura de 30 L y por permitir la utilización del *GEA Niro Soavi homogenizer*.

Al Servicio de Microscopía Electrónica del CIB-CSIC por el apoyo en la preparación de muestra y recogida de datos de tinción negativa y al grupo de Microscopía Electrónica y Reconstrucción Tridimensional de Macromoléculas del CIB por los útiles consejos.

Al Laboratory of Molecular Biology of the Medical Research Council (LMB-MRC) por permitir el acceso a los microscopios electrónicos y por el apoyo en la recogida de datos de crio-microscopía electrónica.

A la Dra. Olga Calvo por la ayuda en el marco de la colaboración con su laboratorio y a todos los autores del artículo relacionado con esta tesis.

A la Dra. Bettina Böttcher por lo que me enseñó en la estancia en su laboratorio.

A mi familia, mi novia y mis amigos.

SUMMARY

In eukaryotes, DNA-dependent RNA polymerase I (Pol I) is the 590 kilodalton enzyme responsible for ribosomal DNA (rDNA) transcription, leading to the synthesis of the ribosomal RNA precursor. Pol I`s activity is crucial for ribosome biogenesis and, therefore, its modulation influences cell growth. The initiation factor Rrn3, conserved in eukaryotes, binds Pol I to form Pol I:Rrn3 complex, which is recruited to the rDNA promoter. Thus, Rrn3-bound Pol I corresponds to an activated state of the enzyme, whose activity can be regulated by assembling or disrupting Pol I:Rrn3. The cryo-electron microscopy (cryo-EM) structures of yeast free monomeric Pol I at 4.9 Å resolution and Pol I in complex with Rrn3 at 7.7 Å are the major findings reported here. The derived pseudo-atomic models unveil the structural changes in the transition from inactive Pol I dimers, previously solved by X-ray crystallography, to free monomers and from these to the activated state bound to Rrn3. In addition, analytical ultracentrifugation suggests that yeast Rrn3 dimers found in solution might establish a dimer-monomer equilibrium *in vitro* upon dilution. Also, electrophoretic mobility shift assays indicate yeast Rrn3 could not bind DNA as previously described for the mammalian homolog. These results provide valuable information on Pol I and Rrn3 changes required for enzyme activation.

PRESENTACIÓN

En los eucariotas, la ARN polimerasa I dependiente del ADN (Pol I) es la enzima de 590 kilodaltons responsable de la transcripción del ADN ribosómico (ADNr), conduciendo a la síntesis del precursor del ARN ribosómico. La actividad de Pol I es crucial para la biogénesis de los ribosomas y, por lo tanto, su modulación influye en el crecimiento celular. El factor de iniciación Rrn3, conservado en eucariotas, se une a Pol I para formar el complejo Pol I:Rrn3, que se recluta al promotor del ADNr. Así, Pol I unida a Rrn3 corresponde a un estado activado de la enzima, cuya actividad puede ser regulada ensamblando o rompiendo Pol I:Rrn3. Las estructuras mediante crio-microscopía electrónica de la Pol I libre y monomérica de levadura a una resolución de 4.9 Å y de la Pol I en complejo con Rrn3 a 7.7 Å son los principales hallazgos aquí descritos. Los modelos pseudo-atómicos resultantes revelan los cambios estructurales en la transición de dímeros inactivos de Pol I, previamente resueltos por cristalografía de rayos X, a monómeros libres y de éstos al estado activado unido a Rrn3. Además, resultados de ultracentrifugación analítica sugieren que los dímeros de Rrn3 de levadura encontrados en solución podrían establecer un equilibrio dímero-monómero *in vitro* tras dilución. Adicionalmente, los ensayos de cambio de movilidad

electroforética indican que la Rrn3 de levadura no se une al ADN, como se describió para la proteína homóloga en mamíferos. Estos resultados proporcionan información valiosa sobre los cambios en Pol I y Rrn3 necesarios para la activación de la enzima.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF VIDEOS

ABBREVIATIONS

Abs_{260nm}	absorbance at 260 nanometres
Abs_{280nm}	absorbance at 280 nanometres
APS	ammonium persulfate
Un	amount of enzyme required to digest 1 µg of pBR322 DNA (dam methylated) in 1 hour at 37°C in a total reaction volume of 50 microliters
U	amount of enzyme that catalyses the incorporation of 10 nanomol of deoxyribonucleotides into a polynucleotide fraction in 30 min at 72°C
Å/pix	Ångström per pixel
cm	centimetre
°C	degree Celsius
$e^-/Å^2$	electrons per Ångström square
g	gram
g	gravitational force
h	hour
kDa	kilodalton
keV	kiloelectronvolt
krpm	kilorevolutions per minute
kV	kilovolt
L	litre
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionization
µg	microgram
µL/min	microliter per minute
µm	micrometre
µmol	micromole
µM	micromolar
mg	milligram
mL	millilitre
mm	millimetre
mM	millimolar

min	minute
M	molar
ng	nanogram
nm	nanometre
nmol	nanomole
OD_{600nm}	optical density at 600 nanometres
p	p value
ppm	parts-per-million
% (v/v)	percentage of a liquid product volume per volume of solution using the same volume units
% (w/v)	percentage of weight of solute in grams per volume of solution in millilitres
pmol	picomole
psi	pound force per square inch
rpm	revolutions per minute
RT	room temperature
s	second
TEMED	tetramethylethylenediamine
TOF	time-of-flight mass spectrometer

INTRODUCTION

1. Transcription

Transcription is a vital biological process that occurs in all cells of all living beings including viruses, as well as in chloroplasts and mitochondria. In a broad sense, transcription is the synthesis of a single-stranded RNA molecule using a specific region of one strand of the DNA as a template. The RNA molecule can contain the information for the synthesis of a particular protein in a process called translation or can directly play a role inside the cell. Transcription can be seen as transference of information from the DNA to the RNA. A cell impeded from transferring the instructions encoded in the DNA would have even the basic functions fatally impaired.

The building blocks of the RNA molecule are ribonucleoside triphosphates, composed of a nitrogenous base (adenine (A), guanine (G), uracil (U) or cytosine (C)), the five-carbon sugar ribose and three phosphate groups. RNA polymerase (RNAP) is the multi-subunit enzyme responsible for transcription. Using ribonucleoside triphosphates, the RNAP extends the RNA chain in the 5′ to 3′ direction by the formation of sequential phosphodiester bonds between the 3′ hydroxyl end of the growing strand and a ribonucleotide, with the release of one pyrophosphate ion.

Two basic reasons explain the participation of auxiliary proteins, termed transcription factors (TFs), in transcription with RNAP: the process is tightly regulated according to cell requirements and the RNAP must be exactly positioned in the beginning of the target gene and accurately finish transcription. Besides proteins, other type of molecules such as RNA or single nucleotides can also modulate transcription.

Specific DNA sequences, globally termed promoter*, TFs and the RNAP establish a network of interactions that participate in DNA opening around the transcription start site (TSS) so the template strand (TS) becomes accessible. This initial stage is called initiation and the macromolecular complex formed by the RNAP, the DNA and the required TFs is termed pre-initiation complex (PIC). After closed complex* opening, RNAP moves along the DNA to catalyse the production of complementary RNA and the initial polymerization of nucleotides is also considered initiation. Then the RNAP escapes the promoter and the RNA chain is extended in coordination with downstream DNA unwinding and transcribed region rewind after RNAP passage. This stage is called elongation and the auxiliary TFs are called elongation factors. They form the elongation complex (EC) together with RNAP, the DNA and the portion of RNA already synthesized. As important as initiation and elongation is termination, the final step in transcription. It stops the EC advance defining the downstream boundary of the gene, the DNA recovers the original double helix structure and the RNAP is released to re-initiate transcription. A pre-

*Words signalled with an asterisk are defined in the glossary.

termination complex is formed, as termination is also tightly regulated and has its own factors. The immature RNA can be processed co-transcriptionally or after EC disassembly.

Bacterial transcription takes place in the cytoplasm and is coupled to translation. The RNA is generally fully functional and does not require processing. Usually the transcription unit contains more than one gene, reason why the RNA codes for more than one protein (polycistronic). Bacteria use a single RNAP to transcribe all genes and transcription initiation does not require the assembly of several TFs in the promoter region.

Eukaryotic transcription takes place in the nucleus. Most often the immature RNA requires processing in order to be exported to the cytoplasm, where it is translated or directly performs some function. Therefore, transcription is not coupled to translation. Also, although alternative splicing* may exist, commonly the translated RNAs code for a single protein (monocistronic). Eukaryotes have three RNAPs specialized in the transcription of distinct subsets of genes and use several TFs to initiate transcription. RNA polymerase I (Pol I) fully dedicates to ribosomal RNA (rRNA) synthesis, RNA polymerase II (Pol II) produces all mRNAs and several snRNAs and RNA polymerase III (Pol III) synthetises small non-coding RNAs (ncRNAs), including all tRNAs and 5S ribosomal RNA (5S rRNA).

Archaeal transcription is considered a hybrid of eukaryotic and bacterial ones. Like in bacteria, it takes place in the cytoplasm and is coupled to translation. The RNA is often polycistronic and does not require processing. All genes are transcribed by a single RNAP. However, the composition and structure of the RNAP, as well as the most important TFs, is more similar to Eukarya.

2. Multi-Subunit RNA Polymerases*

2.1. Bacterial RNA Polymerase

Bacterial RNA polymerase (bRNAP) transcribes all bacterial genes, catalysing the synthesis of mRNA and all ncRNAs. Five subunits (β, β`, two identical α subunits and ω) form the core bRNAP (Figure 1a), which is catalytically active but cannot recognize the promoter and initiate transcription. β` and β are the largest subunits and organize to form a central cleft with two pincers, a characteristic structure dubbed *crab-claw* (Zhang et al 1999) common to all multi-subunit RNAPs (Figure 1b). The cleft allows template DNA access to the active site in the centre of the enzyme, where three conserved aspartic acid residues chelate a Mg^{2+} ion. The two copies of α participate in β`:β assembly and may also interact with the promoter. The smallest subunit, ω, serves as a chaperone for β` (Mathew and Chatterji 2006), facilitating bRNAP assembly and stabilizing the enzyme. The existence of a channel for ribonucleotide substrates entry in the active centre and a second tunnel for RNA-exit were also proposed (Zhang et al 1999). Before starting transcription,

core bRNAP must bind sigma factor (σ) to form the holoenzyme* (Gross et al 1998). σ, a single polypeptide, is sufficient for bacterial transcription initiation (Darst et al 2001, Sekine et al 2012) as it confers promoter specificity.

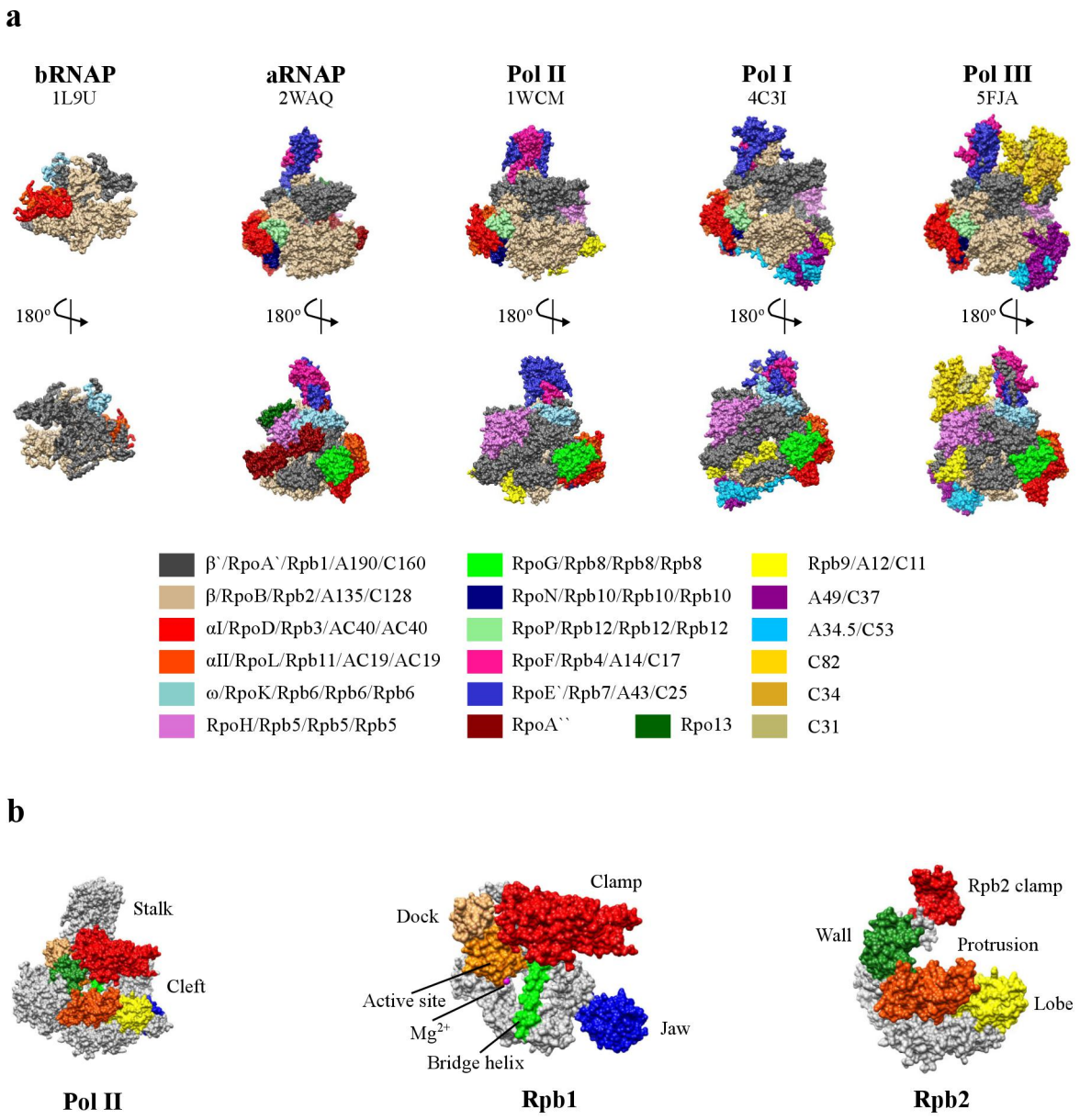

Figure 1. Structures and overall architecture of multi-subunit RNA polymerases.
(a) Atomic models of multi-subunit RNAPs. Conserved subunits are represented in the same colour. RpoA`` shares homology with Rpb1/A190/C160. Rpo13 and RpoG are not conserved in all archaea. PDB codes are indicated. (b) Overall architecture and some universal features of RNAP (Pol II as example, left). Close-up views of Rpb1 (centre) and Rpb2 (right). The stalk is absent in bacteria and the motif correspondent to the dock domain is the β-flap and its tip.

2.2. Archaeal RNA Polymerase

Archaeal RNA polymerase (aRNAP) also transcribes all genes. The 5 subunits forming the core bRNAP are conserved in archaea (and eukaryotes). Depending on the species, the enzyme may consist of 11 to 13 subunits (Figure 1a). X-ray crystal structures of aRNAP revealed its architecture (Hirata et al 2008, Hirata and Murakami 2009, Korkhin et al 2009). Polypeptides A` and A``, encoded from two genes transcribed as one large polycistronic operon (Langer et al 1995), interact to form subunit A`+A`` (orthologous of β`, Table 1). The second largest subunit, B (orthologous of β), contains two $Zn^{2+}$ ions and together with A`+A`` forms a cleft that contains the active site. Comparable to α homodimer in bacteria, subunits D and L form a heterodimer responsible for the assembly and stability of B and A`+ A``. D contains a Fe-S cluster that may play a role in the structural integrity of the subunit. Subunit K (orthologous of ω) helps connecting A`+A`` with the stalk, a sub-complex that protrudes from aRNAP, absent in bRNAPs, formed by subunits E` and F. Subunit H is placed around the same region of A`` that binds K. Subunits N and P bind B, occupy small clefts and do not protrude from the enzyme. Subunit G, absent in some species, interacts with A` and sits in an external crevice formed by it. Rpo13, which is only found in the order *Sulfolobales*, is located at a groove between H and A` and it was suggested to play a role in the formation of the transcription bubble* based on its location (Korkhin et al 2009). Subunits E`, F, H, N, P, G and Rpo13 are not conserved in bacteria, although some regions of bacterial subunits may occupy similar positions and have equivalent functions.

2.3. Eukaryotic RNA Polymerases

2.3.1. RNA Polymerase II

RNA polymerase II (Pol II) is a 12-subunit macromolecular enzyme (Figure 1a) that produces all eukaryotic mRNAs and several snRNAs. Rpb1 and Rpb2, like their counterparts in bacteria and archaea (Table 1), form the cleft and the active centre with the Mg^{2+} ion. Rpb3 and Rpb11 form a heterodimer and function as the α/α or D/L sub-complexes. There are 5 subunits common to the three eukaryotic RNAPs: Rpb5, Rpb6, Rpb8, Rpb10, Rpb12. All 5 have homologs in archaea, being Rpb6 homologous of bacterial ω (Minakhin et al 2001), and are especially important to maintain the macromolecular complex architecture and stability. Rpb9, without homolog in Archaea, contacts the lobe and the jaw through interactions with Rpb2 and Rpb1, respectively. It was suggested a role in transcription fidelity for Rpb9 (Walmacq et al 2009). Rpb4/Rpb7 heterodimer forms a stalk as the sub-complex E`/F in archaea.

Table 1. Homology in RNAP subunits and initiation factors.
Initiation factors are in blue and TFS/TFIIS in orange. Rrn7 and TFIIB share sequence and functional homology but lack structural homology (Engel et al 2017).

bRNAP	aRNAP	Pol II	Pol I	Pol III
β`	RpoA`+RpoA`` (Rpo1)	Rpb1	A190	C160
β	RpoB (Rpo2)	Rpb2	A135	C128
α	RpoD (Rpo3)	Rpb3	AC40	AC40
α	RpoL (Rpo11)	Rpb11	AC19	AC19
ω	RpoK (Rpo6)	Rpb6 (ABC23)	Rpb6 (ABC23)	Rpb6 (ABC23)
	RpoH (Rpo5)	Rpb5 (ABC27)	Rpb5 (ABC27)	Rpb5 (ABC27)
	RpoG (Rpo8)	Rpb8 (ABC14.5)	Rpb8 (ABC14.5)	Rpb8 (ABC14.5)
	RpoN (Rpo10)	Rpb10 (ABC10β)	Rpb10 (ABC10β)	Rpb10 (ABC10β)
	RpoP (Rpo12)	Rpb12 (ABC10α)	Rpb12 (ABC10α)	Rpb12 (ABC10α)
		Rpb9	A12.2 N-ter	C11 N-ter
	TFS	TFIIS	A12.2 C-ter	C11 C-ter
	RpoF (Rpo4)	Rpb4	A14	C17
	RpoE` (Rpo7)	Rpb7	A43	C25
	Rpo13			
		TFIIFα (Tfg1)	A49 N-ter	C37
		TFIIFβ (Tfg2)	A34.5	C53
	TFE	TFIIEβ (Tfa2)	A49 C-ter	C34
				C82
				C31
	TBP	TBP	TBP	TBP
	TFB	TFIIB	Rrn7	Brf1

2.3.2. RNA Polymerase III

RNA polymerase III (Pol III), the largest among all RNAPs, consists of 17 different subunits (Figure 1a) and produces small ncRNAs including all tRNAs and 5S rRNA. Apo-Pol III cryo-EM structure at sub-nanometric resolution was recently determined (Hoffmann et al 2015). C160 and C128, with homologs in all multi-subunit RNAPs (Table 1), form the active centre and the cleft. The heterodimer AC40/AC19 is shared with Pol I and conserved in the other multi-subunit RNAPs, while 5 subunits are common to all eukaryotic RNAPs and have homologs in archaea (Rpb5, Rpb6, Rpb8, Rpb10, Rpb12). The N-terminus (N-ter) of C11 is homologous to Rpb9. C25/C17 sub-complex forms the conserved stalk in archaea and eukaryotes. In addition to these 12 subunits with counterparts in archaea and eukaryotes, two additional sub-complexes decorate Pol III. C37/C53

heterodimer bears topology, structural and functional similarities with A49/A34.5 in Pol I. C31/C82/C34 heterotrimer is Pol III exclusive.

2.3.3. RNA Polymerase I

RNA polymerase I (Pol I) is a 14-subunit macromolecular enzyme (Figure 1a) fully dedicated to the synthesis of rRNA, major component of the ribosome. Apo-Pol I crystal structures corresponding to the inactive dimeric form of the enzyme were solved (Engel et al 2013, Fernández-Tornero et al 2013). A190 and A135, the two largest subunits that constitute the DNA-binding cleft and the active site, are homologous of the cleft-forming subunits from bacteria, archaea and eukaryotic multi-subunit RNAPs (Table 1). AC40 and AC19 form a heterodimer critical for cleft assembly and are homologous of the corresponding dimers in bacteria, archaea and Pol II. Five subunits, Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12, are shared with the other eukaryotic RNAPs and have homologous in archaea. A12.2 N-ter is homologous to Rpb9 and A12.2 C-terminus (C-ter) can remodel the active centre of the enzyme to confer RNA cleavage activity. A14 and A43 assemble a heterodimer protruding from Pol I, the stalk, homologous to archaeal and eukaryotic stalks. The 12 mentioned subunits are completed with two extra proteins, A49 and A34.5, forming a heterodimer at the periphery, protruding opposite the stalk. A49 C-ter forms a tandem winged helix (A49tWH) (Geiger et al 2010) not solved in Pol I dimeric crystal structures.

2.4. A Common Ancestor for Multi-Subunit RNA Polymerases

Multi-subunit RNAPs from the three domains of life share structural homology (Hirata et al 2008, Ruprich-Robert and Thuriaux 2010, Werner and Grohmann 2011). Their two biggest subunits (archaeal A`+A`` considered one single subunit) form a DNA-binding cleft that harbours the active centre in the wall and adopt a *crab-claw* shape whereas two conserved additional subunits always constitute a minimal assembly platform. Multi-subunit RNAPs have three conserved aspartic acid residues in the largest subunit (Sosunov et al 2003) that coordinate at least one divalent metal ion in the catalytic centre. Also, all 5 subunits of core bRNAP have homologous subunits in Archaea and Eukarya (Table 1), strongly suggesting a common ancestor.

Evolution from bRNAP to aRNAP brought changes in architecture, number of subunits and size. aRNAP displays 6 to 7 extra subunits and the RpoF/RpoE` subcomplex forms the stalk, a recognizable feature from archaeal and eukaryotic RNAPs and most pronounced difference to bRNAP (Werner and Grohmann 2011). The stalk is a versatile region implicated in initiation, elongation and termination (Werner and Grohmann 2011), dissociable in Pol II, at least in *S. cerevisiae* (Orlicky et al 2001), but tightly bound in aRNAP (Grohmann et al 2009), Pol I

(Fernández-Tornero et al 2013) and Pol III (Hoffmann et al 2015). aRNAP resembles much more eukaryotic Pol II than bRNAP (Hirata et al 2008, Korkhin et al 2009). The overall structure of aRNAP and Pol II resembles a *crab claw* with a protruding stalk. The relative positioning of the stalk in relation to the central subunits is highly conserved between aRNAP (Hirata and Murakami 2009) and the three nuclear RNAPs (Fernández-Tornero et al 2007), suggesting the heterodimer was acquired by a common ancestor RNAP (Hirata and Murakami 2009). Concerning the two biggest subunits, structural similarity between Rpb1 in Pol II and A`/A`` in aRNAP was estimated to be 68.9% whereas between Rpb1 and β` in bRNAP only 50% (Hirata et al 2008). Rpb2 and B are 90% alike while Rpb2 and β share only 64.9% structural similarity. Also, the folding of highly conserved segments around the active site is essentially identical in Pol II and aRNAP (Hirata and Murakami 2009), most differences between the enzymes being greater as the distance to the central part increases. Furthermore, almost all the structural discrepancies between archaeal and eukaryotic RNAPs can be classified as simple addition of Pol II-specific polypeptides (Rpb9, Rpb5 jaw and Rpb1 C-terminal domain (CTD)) to aRNAP rather than changes to the main RNAP architecture. All aRNAP subunits have homologs in Pol II and the only exception is Rpo13, found in some but not all archaeal species. Accordingly, only Pol II subunit Rpb9 is not conserved in aRNAP. During evolution, it seems likely that the structure of aRNAP has been maintained from the ancestor of archaea and eukaryotes whereas nuclear RNAPs further accompanied the evolutionary lines of the Eukarya.

2.5. Pol I and Pol III Incorporated "Built-in" Transcription Factors

Concerning eukaryotic multi-subunit RNAPs, all 12 Pol II subunits are either shared or homologous to Pol I or Pol III subunits. Interestingly, Pol I and Pol III incorporated additional subunits or domains with functional similarity to some Pol II TFs (Table 1).

A12.2 in Pol I and C11 in Pol III, homologous of Rpb9 in their N-terminus, have C-terminal extensions homologous to Pol II TFIIS (Vannini and Cramer 2012). In the recovery from transcription arrest (result of a DNA mutation, for instance), during transcriptional proofreading* or in other situations, multi-subunit RNAPs "backtrack" and cleave the nascent RNA chain to resume transcription. In Pol I and Pol III, the RNA cleavage is intrinsic to the RNAP and depends on a specific subunit, A12.2 or C11, respectively, that reaches the active site and can change its catalytic activity from RNA polymerization to its reverse. In turn, Pol II intrinsic pyrophosphorolysis activity is weak and the enzyme requires TFIIS to cleave the RNA (Gu and Reines 1995).

TFIIF is a general transcription factor* (GTF) part of Pol II PIC. TFIIF roles in initiation include overall PIC stabilization (Tan et al 1994), influence in TSS selection (Ghazy et al 2004), stimulation of phosphodiester bond formation and early RNA synthesis (Ren et al 1999, Yan et al

1999, Khaperskyy et al 2008). Sub-complexes A49/A34.5 in Pol I, C53/C37 in Pol III and Pol II TFIIF have a structurally conserved dimerization module (Carter and Drouin 2010, Geiger et al 2010) that binds the polymerase lobe similarly (Vannini and Cramer 2012). A49/A34.5 dimerization domain stimulates RNA cleavage (Geiger et al 2010), most likely through the stabilization of the A12.2 N-ter and, therefore, influencing the position of A12.2 catalytic C-ter in the cleft (Engel et al 2013, Fernández-Tornero et al 2013). In Pol II, a stimulation of TFIIS RNA cleavage by TFIIF might also occur, as TFIIF suppresses Pol II pausing in Drosophila (Price et al 1989).

A49 subunit in Pol I contains a C-ter tandem winged-helix domain (A49tWH), a DNA-binding motif absent in TFIIF but predicted in TFIIE (Geiger et al 2010), other GTF component of Pol II PIC. TFIIE binds single-stranded DNA (Okamoto et al 1998) and double-stranded DNA (Okuda et al 2000, Tanaka et al 2009), similarly to A49tWH (Geiger et al 2010). Furthermore, in a Pol I transcribing complex, A49tWH contacts upstream DNA similar to TFIIE in Pol II initiation complexes (Tafur et al 2016). Thus, distinct data point to a structural and functional relation between A49tWH and TFIIE. Among other roles in initiation, TFIIE is able to open some promoters and stabilize the open complex* (OC) (Holstege et al 1995, Holstege et al 1996, He et al 2013, Plaschka et al 2016). Curiously, A49tWH was recently shown to support initiation *in vitro* (Pilsl et al 2016 a). However, A49/A34.5 and A49tWH domain were also shown to stimulate Pol I elongation (Kuhn et al 2007, Beckouet et al 2008, Geiger et al 2010) while TFIIE role seems limited to Pol II initiation. Interestingly, neither C53 subunit from Pol III, homologous to A49, contains a tWH nor Pol III operates with an independent TFIIE-related factor. Instead, subunit C34 from Pol III C82/C34/C31 specific heterotrimer presents winged-helix domains and homology in some regions to TFIIE (Carter and Drouin 2010, Vannini and Cramer 2012, Hoffmann et al 2015).

Pol II targets a heterogeneous set of genes and requires further levels of regulation, thus having a more diverse and complex transcription machinery. On the other hand, the genes targeted by Pol I and Pol III are less heterogeneous (Pol I fully dedicated to the transcription of a single type of gene) but require ready and efficient transcription to meet cellular needs during growth, ribosome biogenesis and translation. The additional sub-complexes found in Pol I and Pol III work as "built-in" TFs that allow efficient transcription of Pol I and Pol III DNA targets.

2.6. Unique Features of Multi-Subunit RNA Polymerases

Core bRNAP is not able to initiate transcription, as it must bind σ to form the holoenzyme and recognize the promoter. The bRNAP:σ complex is able to initiate transcription without additional TFs and this constitutes a unique bRNAP feature. Bacteria possess a housekeeping σ that recognizes the promoters controlling the transcription of most genes and a variable number of

alternative σs with different promoter-recognition properties (Gruber and Gross 2003). For instance, *E. coli* σ^{70} recognizes most promoters while 6 other σs are expressed in response to specific stresses (Colland et al 2002). Apart from σ^{54} family, all σs share common features (Merrick 1993). They are multi-domain proteins composed of up to 4 domains joined by linkers (Campbell et al 2002).

In Pol II, Rpb1 CTD is unstructured, consists of repeats of 7 amino acids (YSPTSPS) that range from 26 in yeast to 52 in vertebrates and it is target of kinases and phosphatases (Jeronimo et al 2016). The CTD is unique of Pol II and allows modulation of initiation, elongation, termination, capping* of mRNA precursor and splicing (Hsin and Manley 2012). The CTD reflects the need for regulation of Pol II compared to any other multi-subunit RNAP.

A unique feature of Pol III, contrary to the other multi-subunit RNAPs and likely related to the short genes Pol III transcribes, is that Pol III PIC assembly requires interactions with internal promoter elements in addition to the binding of DNA sequences located upstream and around the TSS.

Pol I ability to dimerize into transcriptional inactive dimers through an interaction between the C-ter of A43 and the cleft of the neighbouring enzyme (Engel et al 2013, Fernández-Tornero et al 2013, Kostrewa et al 2015) also seems to be a distinctive feature (Torreira et al 2017). The dimerization was observed in solution (Milkereit et al 1997, Bischler et al 2002) and in living yeast cells (Torreira et al 2017). In addition, Pol I is unique in its recruitment to the rDNA promoter. It requires the binding of the Rrn3 initiation factor like bRNAP needs σ association. However, while bRNAP holoenzyme can initiate transcription, Pol I:Rrn3 complex further requires promoter-bound initiation factors. aRNAP, Pol II and Pol III do not rely on a specific recruitment factor, as all three can bind the growing PIC by themselves.

3. Pol I Transcription Initiation

3.1. rDNA

The nucleolus forms around the nucleolar organiser regions (NORs). In humans, NORs are located in the short arms of acrocentric chromosomes* and in yeast there is a single NOR in chromosome XII. NORs contain several tandem repeats of ribosomal DNA (rDNA) genes, approximately 150-200 in yeast and 400 in human diploid cells (Birch and Zomerdijk 2008). In yeast, each repeat has in total approximately 9.1 kb and consists of the small 5S rRNA gene, a 6.9 kb transcribed region coding for the 35S pre-rRNA and non-transcribed regions including Pol I promoter (Figure 2a). In mammals, each rDNA repeat is approximately 43 kb, the non-transcribed regions around 30 kb and the transcribed region coding for the immature 47S pre-rRNA approximately 13 kb (McStay and

Grummt 2008). The general layout of the rDNA promoter is conserved from yeast to humans, with similar spacing between TSS and promoter elements, but there is little sequence similarity (Goodfellow and Zomerdijk 2013). Promoters of eukaryotic rDNA genes contain two main regions: the upstream activating sequence (UAS) and core element (CE) in yeast (Figure 2a) and the upstream core element (UCE) and CE in humans. In both species the CE overlaps the TSS and the UAS/UCE is centred approximately 100 bps upstream the TSS (Drygin et al 2010). At least in yeast the CE is essential and the UAS, although not strictly required, is important for a high level of transcription (Keys et al 1996).

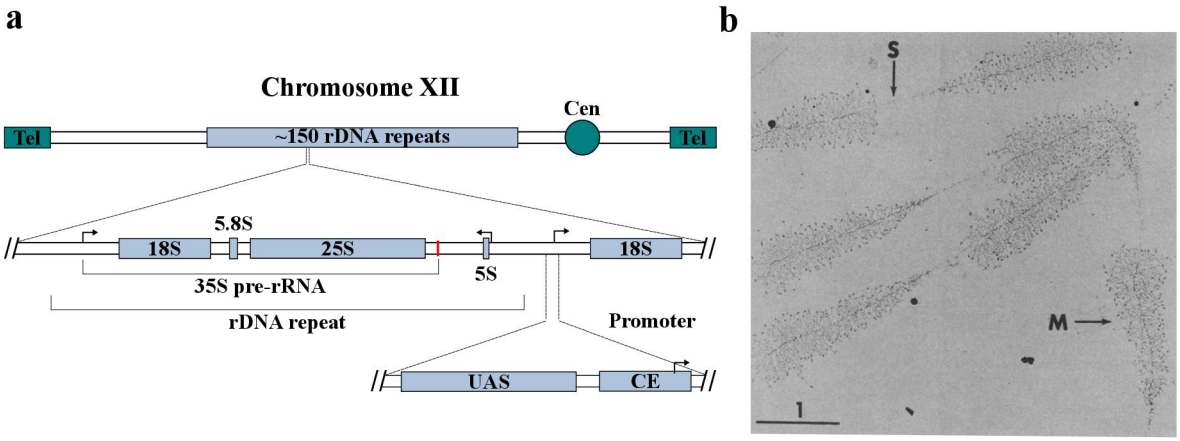

Figure 2. rDNA organization.
(a) Schematic of *S. cerevisiae* rDNA locus. Tel, telomer; Cen, centromer; UAS, upstream activating sequence; CE, core element; red line, end of Pol I-transcribed DNA; arrows, TSSs. (b) Miller spread of a *Triturus viridescens* oocyte (Miller and Beatty 1969). The bar corresponds to 1 μm. S, non-transcribed spacer; M, transcribed matrix.

The rDNA repeats being transcribed by Pol I can be observed by electron microscopy and form structures called Miller spreads resembling Christmas trees (Miller and Beatty 1969) (Figure 2b). Intercalated transcribed and non-transcribed regions are visible, as well as several Pol I molecules loaded in a single rDNA gene (black dots in the "tree trunk"). In fact, Pol I high loading rate seems to be a distinctive feature of the enzyme, as in exponentially growing cells 0.01 to 10 Pol II enzymes are found per open reading frame (Bon et al 2006) whereas more than 120 Pol I molecules can be detected in Miller spreads at each active rDNA repeat (Osheim et al 2009). Miller spreads also exhibit pre-rRNA molecules being synthesised, corresponding to the "branches" that increase in length over the gene. At the end of the "branches" there are "knobs" where the immature RNA is co-transcriptionally processed (Mougey et al 1993).

3.2. Pol I Transcription Initiation in Yeast

In yeast, Pol I PIC assembly requires 4 GTFs: upstream activating factor (UAF), TATA-box* binding protein (TBP), core factor (CF) and Rrn3 (Figure 3). The most accepted model for Pol I PIC assembly involves the binding of UAF to UAS, TBP attachment followed by CF interaction with CE and subsequent Pol I recruitment in complex with Rrn3. UAF, CF and Rrn3 were found only in the promoter (Bier et al 2004) and the same should be expected from TBP. It is thought these 4 initiation factors do not participate in elongation.

UAF is composed of 6 subunits (Keys et al 1996, Keener et al 1997, Siddiqi et al 2001) and is poorly conserved in higher eukaryotes. UAF forms a stable complex with the UAS in the promoter (Keys et al 1996, Keener et al 1997) and seems not to require CF, Pol I or Rrn3 for the interaction (Goetze et al 2010). Rrn9 subunit of UAF interacts strongly with TBP *in vitro* and in a yeast two-hybrid assay (Steffan et al 1996) and this interaction is important for high-level transcription of rDNA by Pol I *in vivo* (Steffan et al 1998). Besides its role as a PIC component, UAF participates in chromatin architecture rearrangements of the entire Pol I-transcribed region (Goetze et al 2010) and this is related to the inhibition of rDNA genes transcription by Pol II (Vu et al 1999, Goetze et al 2010).

TBP is also required as a PIC component although Pol I promoter does not contain a TATA-like element. In addition to UAF, TBP binds the CF (Lin et al 1996). A specific interaction between TBP and Rrn6 subunit of CF was found *in vitro* and by yeast two-hybrid, suggesting a role for TBP in the recruitment of the CF if UAF is present (Steffan et al 1996). The exact role of TBP in Pol I transcription is not yet clear, but UAF, TBP and CF probably cooperate to recognize the promoter *in vivo*.

The three-subunit CF complex (Lalo et al 1996, Lin et al 1996) plays a central role in PIC assembly (Figure 3). It binds the promoter CE and TBP, but also Rrn9 subunit of the UAF *in vitro* and in a yeast two hybrid assay (Steffan et al 1996). Moreover, *in vitro*, the CF contacts Pol I and Rrn3 in Pol I:Rrn3:CF complex and in a Pol I initially transcribing complex* (ITC) without UAF and TBP (Engel et al 2017). Interestingly, the CF does not strictly require UAF or TBP for promoter recognition, as transcription can be initiated *in vitro* in a minimal promoter only in the presence of CF, Pol I and Rrn3 (Bedwell et al 2012, Pilsl et al 2016 b). However, separately adding UAF or TBP further activates Pol I transcription (Bedwell et al 2012).

Rrn3, functionally conserved in eukaryotes from yeast to human (Schnapp et al 1993, Bodem et al 2000, Moorefield et al 2000), was found to interact with subunit A43 of Pol I (Peyroche et al 2000). Once UAF, TBP and CF form a stable complex with the promoter (or CF alone), Rrn3-bound Pol I binds to the already assembled PIC (Knutson and Hahn 2013). Rrn3 interacts with Pol I before recruitment of the enzyme to rDNA promoter. Pol I:Rrn3 complex is the activated form of

the enzyme, ready for promoter recognition and necessary to initiate transcription. In the absence of Rrn3, Pol I can still be recruited but does not yield a productive PIC (Aprikian et al 2001). Rrn3 was also proposed to bind Rrn6 subunit of CF in its C-ter, based on a yeast two hybrid assay and co-elution of Rrn3 and Rrn6 C-ter from affinity chromatography columns *in vitro* (Peyroche et al 2000). Such physical interaction was not confirmed in Pol I:Rrn3:CF or Pol I ITC structures (Engel et al 2017). The single interaction observed between Rrn3 and CF in those macromolecular assemblies involved Rrn7 but not Rrn6.

Rrn3 leaves the transcribing complex during promoter escape or very early in elongation and for sure is not part of a stable Pol I EC (Milkereit and Tschochner 1998, Bier et al 2004). *In vitro*, UAF remains stably bound to the promoter whereas CF, TBP and Rrn3 are released upon initiation, suggesting that UAF may operate as a re-initiation scaffold for subsequent rounds of transcription (Aprikian et al 2001). However, *in vivo*, UAF, TBP and CF keep stably associated to the promoter in the absence of Rrn3 and Pol I, possibly implying that CF does not leave the promoter with Pol I and Rrn3 and belongs to the re-initiation scaffold (Goetze et al 2010). It was also proposed that, *in vivo*, CF recognition of the promoter CE depended on the binding of Pol I and Rrn3 (Bordi et al 2001), but those data are difficult to conciliate with the current knowledge on Pol I PIC assembly.

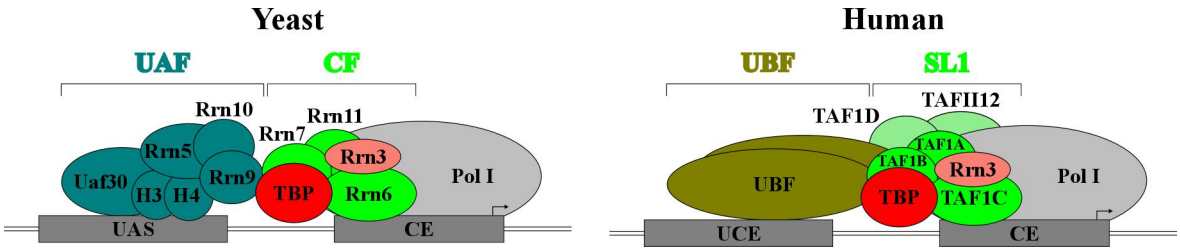

Figure 3. Pol I pre-initiation complexes in yeast and human.
Adapted from Knutson and Hahn 2013. UBF and subunits TAF1D and TAFII12 from SL1 are coloured differently because they are not conserved. TBP is part of SL1 but not of the CF.

3.3. Pol I Transcription Initiation in Higher Eukaryotes

In higher eukaryotes, Pol I PIC formation requires at least three GTFs: upstream binding factor (UBF), not sharing homology with none of the yeast UAF subunits (Claypool et al 2004, Knutson and Hahn 2013), human selective factor 1 (SL1) (named TIF-IB in mouse), with some similarity to CF and containing TBP, and Rrn3 (also known as TIF-IA) (Russell and Zomerdijk 2005) (Figure 3). Pol I PIC formation involves the sequential binding of UBF to promoter UCE, SL1 association to promoter CE and, similarly to yeast, final Pol I:Rrn3 recruitment.

UBF is a dimer (McStay et al 1991) and as a dimer binds the promoter and induces substantial topological changes (Jantzen et al 1990, Bazett-Jones et al 1994). UBF forms a loop that bridges

the UCE and the CE (Bazett-Jones et al 1994, Stefanovsky et al 2001) and can trigger large-scale chromatin decondensation (Chen et al 2004). It was suggested that UBF binds the promoter and the entire transcribed region (O`Sullivan et al 2002, Sanij and Hannan 2009), plays its role after PIC formation and before elongation, at promoter escape (Panov et al 2006), and also regulates elongation taking advantage of its interactions with the transcribed gene (Stefanovsky et al 2006).

SL1 is a 6-subunit complex (Denissov et al 2007), three subunits orthologous of yeast CF subunits (Boukhgalter et al 2002, Russell and Zomerdijk 2006) and TBP among the other three polypeptides. SL1 was shown to bind the promoter (Rudloff et al 1994, Beckmann et al 1995), Rrn3 (Miller et al 2001), UBF (Beckmann et al 1995) and, in mouse, TIF-IB binds Pol I subunit PAF49 (Hanada et al 1996, Yamamoto et al 2004), homolog of A34.5 in yeast. Although it was initially proposed that binding of UBF to the UCE leaded to the recruitment of the TBP-containing complex SL1 (Learned et al 1986, Bell et al 1988), later it was shown that SL1 can bind the promoter in the absence of UBF and direct initiation (Friedrich et al 2005), analogous to CF in yeast.

UBF present or not, TIF-IB/SL1 contacts promoter DNA and probably recruits an initiation-competent Pol I subpopulation complexed with TIF-IA/Rrn3. TIF-IA/Rrn3 behaves similarly to yeast Rrn3 in Pol I initiation. Only the subpopulation of Pol I associated to TIF-IA/Rrn3 supports transcription initiation (Miller et al 2001). Human Rrn3 physically interacts with two SL1 subunits and blocking this interaction prevents Pol I:Rrn3 recruitment to the promoter (Miller et al 2001). In mouse, a short conserved motif was identified as the responsible for the association of TIF-IA with two TIF-IB subunits and, in the context of the rDNA promoter, interactions between TIF-IA and two Pol I subunits were identified, being one of those subunits A43 as it happens in yeast (Yuan et al 2002). TIF-IA is liberated from the transcribing complex before a stable EC is formed (Schnapp et al 1993).

Not described to date in yeast is the ability of mammalian Rrn3 to bind DNA (Stepanchick et al 2013). Deletion of the Rrn3 DNA-binding domain abolishes the interaction with the DNA and impairs Pol I transcription *in vitro*, while not affecting the interactions with SL1 and A43 in Pol I. DNA-binding capacity of mammalian Rrn3 was proposed to be essential for Pol I transcription (Stepanchick et al 2013).

4. Regulation of Pol I Transcription

Pol I produces the 35S pre-rRNA, precursor of the major rRNA components of the ribosome. Therefore, Pol I transcription is decisive for the synthesis of ribosomes. Ribosome biogenesis represents a major metabolic effort for the cell. It is amazing how a few hundred rDNA genes are able to produce up to 80% of the total cellular RNA (Moss et al 2007) but fully justified, given that

ribosomal synthesis is paramount in the control of cellular and even organism growth (Moss et al 2007). Ribosome biogenesis demands the coordinated effort of all three nuclear RNAPs (Laferté et al 2006). Pol I synthesises rRNA, Pol III the 5S rRNA and Pol II the mRNAs coding for the ribosomal proteins and factors required for Pol I and Pol III activity and ribosomal organization. In a rapidly growing yeast cell, 60% of total transcription is devoted to rRNA whereas 50% of Pol II transcription and 90% of mRNA splicing are dedicated to ribosomal proteins (Warner 1999). Thus, rDNA transcription must be tightly controlled in response to cell environment changes. Levels of growth factors, amino acids or glucose, as well as cellular stresses, can modulate rRNA synthesis.

Not surprisingly, rDNA transcription is regulated in different ways. At least in mammals, rDNA repeats are not identical copies (Tseng et al 2008, Santoro et al 2010), as they can be differentially regulated in distinct tissues (Tseng et al 2008) and intergenic transcripts from specific subclasses of rDNA repeats can induce methylation and heterochromatin formation to silence other rDNA repeats (Santoro et al 2010). Furthermore, only a subset of the repeats are transcribed in a given period: inactive rDNA genes are tightly packaged into chromatin, DNA in their coding region is methylated, they lack Pol I and UBF and are not at the nucleolus centre, whereas active rDNA genes are less packaged and methylated, associate with Pol I and UBF and are placed in the interior of the nucleolus (Németh and Längst 2011). Finally, transcription can be regulated in initiation, elongation and termination. Concerning initiation, each step in the PIC assembly constitutes an opportunity for regulation of rDNA transcription in yeast and higher eukaryotes. In general, all stimuli blocking or relenting ribosome biogenesis and, therefore, rDNA transcription, have the potential to affect PIC formation. A signal transduction pathway is often the link between an external or internal environmental stimulus and the post-translational modification of a given GTF or Pol I. UBF, SL1, Rrn3 and Pol I are targets of signalling cascades that culminate with specific phosphorylation, activating or repressing transcription (Kusnadi et al 2015). Also important is the TF Myc, involved in the regulation of Pol I transcription through direct interactions with Pol I PIC components (Kusnadi et al 2015).

4.1. Rrn3, Pol I and Pol I:Rrn3 as Targets for the Regulation of Pol I Transcription Initiation in Yeast

In yeast, only a small fraction of Rrn3 in total cell lysates is associated to the more abundant Pol I enzyme (Milkereit and Tschochner 1998, Bier et al 2004) and Rrn3 is in greater concentration than UAF and CF (Bier et al 2004). In addition, under rapamycin repression of Pol I transcription, a strain with Pol I constitutively bound to Rrn3 leaded to abnormal accumulation of all ribosomal components, including 5S rRNA produced by Pol III and ribosomal proteins whose mRNA results from Pol II activity (Laferté et al 2006). These results suggest a regulatory role for Pol I:Rrn3 in

rDNA transcription and how this last biological event influences ribosome synthesis. The first evidence of Pol I:Rrn3 regulatory role came with the discovery that phosphorylation of Pol I is required to form a stable and transcriptionally active Pol I:Rrn3 complex, leading to the hypothesis that phosphorylation/dephosphorylation of Pol I might modulate rDNA transcription (Fath et al 2001). A study identified phosphosites in yeast Pol I but none of the single mutations on those sites revealed the residue critical for Pol I:Rrn3 interaction, possibly because a combination of phophorylations is required or the crucial phosphosite could not be identified (Gerber et al 2008). Besides, inhibition of the TOR* signalling system with rapamycin in yeast cells caused a decrease in the amount of Pol I:Rrn3 complexes (Claypool et al 2004), and although the exact mechanism remains unknown, it implies the TOR pathway regulates the levels of Pol I:Rrn3 inside the cell and, therefore, Pol I recruitment to the promoter.

Furthermore, the 9-subunit Ccr4-Not complex, with roles on Pol II transcription regulation, mRNA decay and quality control, RNA export, translational repression and protein ubiquitination (Miller and Reese 2012), was recently shown to bridge mTORC1 (mechanistic target of rapamycin complex 1) to Pol I regulation (Laribee et al 2015). Although other players must intervene, upon mTORC1 inhibition, Ccr4-Not is required to decrease Pol I transcription as its disruption activates the process (Laribee et al 2015). Also, under nutrient rich conditions, Ccr4-Not suppresses Pol I initiation and Pol I:Rrn3 complex formation is affected. Moreover, Ccr4-Not directly associates with Pol I in a co-immunoprecipitation assay. The data support Ccr4-Not as a Pol I transcriptional regulator that controls Pol I initiation. Evidences for a Ccr4-Not role on the regulation of Pol I elongation were also presented in the same study.

There are no data defining a specific Rrn3 phosphorylation required for *in vivo* interaction with Pol I or impeding Pol I:Rrn3 formation. Yeast Rrn3 is found predominantly phosphorylated *in vivo* when not bound to Pol I and unphosphorylated Rrn3 is able to bind Pol I and transcribe *in vitro* (Fath et al 2001). Accordingly, Rrn3 displays 8 serine residues clustered in 4 pairs in its N-ter surface and a phospho-mimetic mutation in one of those residues (S145D) caused a severe growth phenotype and impaired Pol I:Rrn3 interaction *in vitro* (Blattner et al 2011). Besides, in the S145D Rrn3 mutant strain, Pol I association to rDNA gene, as well as Rrn3 presence in the promoter and at the beginning of the transcribed region, are significantly reduced. Thus, S145 in Rrn3 is involved in Pol I binding and its phosphorylation, not confirmed *in vivo*, could possibly regulate Pol I recruitment to the promoter. Other phospho-mimetic mutation in one of the aforementioned serine residues (S185D) strongly affected the formation of Pol I:Rrn3 *in vitro* (Blattner et al 2011).

Spt6 is an essential protein in yeast, conserved in humans and known to regulate Pol II initiation, elongation and termination. However, only recently it was shown to physically associate with A43 subunit of Pol I (Engel et al 2015). In addition, a mutation in Spt6 affected Pol I occupancy of the rDNA, rRNA synthesis and PIC assembly (Engel et al 2015). Spt6 seems to influence Pol I

transcription independently of other known initiation factors, as Spt6 inactivation reduced Rrn3 and Pol I:Rrn3 levels but overexpression of Rrn3 rescued Pol I:Rrn3 complex formation but not rRNA synthesis (Engel et al 2015). These data suggest Spt6 is involved in either recruiting Pol I:Rrn3 complex or stabilizing the PIC, therefore regulating rRNA synthesis in initiation.

Two poorly understood mechanisms possibly regulate Rrn3 levels in the cell without post-translational modifications. Yeast Rrn3 promoter is targeted by Met32, which in turn recruits the co-activator of yeast sulphur metabolism Met4 (Carrillo et al 2012), suggesting Met4 and Met32 bind to regulate Rrn3 gene transcription. Also, deletion of Sfp1, protein involved in the glucose-dependent regulation of ribosome biogenesis, leads to lower levels of Rrn3 and some Pol I subunits (Cipollina et al 2008), what might indicate that Sfp1 is a TF important for the adequate transcription of Rrn3 and of those Pol I subunits.

Finally, both Pol I and Rrn3 were found as dimers in solution (Milkereit et al 1997, Bischler et al 2002, Blattner et al 2011). Pol I homodimerization in living yeast cells and a role for Pol I dimers in the regulation of rDNA transcription were shown (Torreira et al 2017). Rrn3 dimer formation *in vivo* remains to be confirmed, as well as if the activating factor dimerization constitutes a mechanism of Pol I transcription modulation.

5. Prior Knowledge on Pol I:Rrn3 Architecture

X-ray crystal structures of yeast Pol I (Engel et al 2013, Fernández-Tornero et al 2013) were nearly identical and yielded Pol I in a transcriptionally inactive state characterised by dimerization, a DNA-mimicking loop preventing downstream DNA accommodation near the active site and a wide open cleft unable to trap the DNA. These features may be used to regulate the activity of the enzyme, as active Pol I requires disruption of the dimers, reorganization of the DNA-mimicking loop and cleft closure.

In Pol I crystals, the most extensive interaction in the dimers involves A43 C-ter (251-326) insertion in the cleft of the neighbouring enzyme (Figures 4a and b). A43 C-ter folds into a long α-helix (273-293) followed by a β-hairpin (298-309). The acidic C-ter tail (311-326) was only partially solved (311-316) but also binds the neighbouring Pol I cleft. Actually, a mutation in A43 partially truncating the β-hairpin and deleting the acidic C-ter tail (A43Δ307-326) was enough to disrupt Pol I dimers formation *in vivo* (Torreira et al 2017).

The DNA-mimicking loop in A190 is a Pol I-specific insertion that contains three short helices (Fernández-Tornero et al 2013) and extends from the inner face of A135 lobe to the internal surface of the clamp on the opposite side of the cleft, therefore occupying the TS and RNA sites in a transcribing complex (Figure 4b). In Pol I inactive dimers, the DNA-mimicking loop binds the

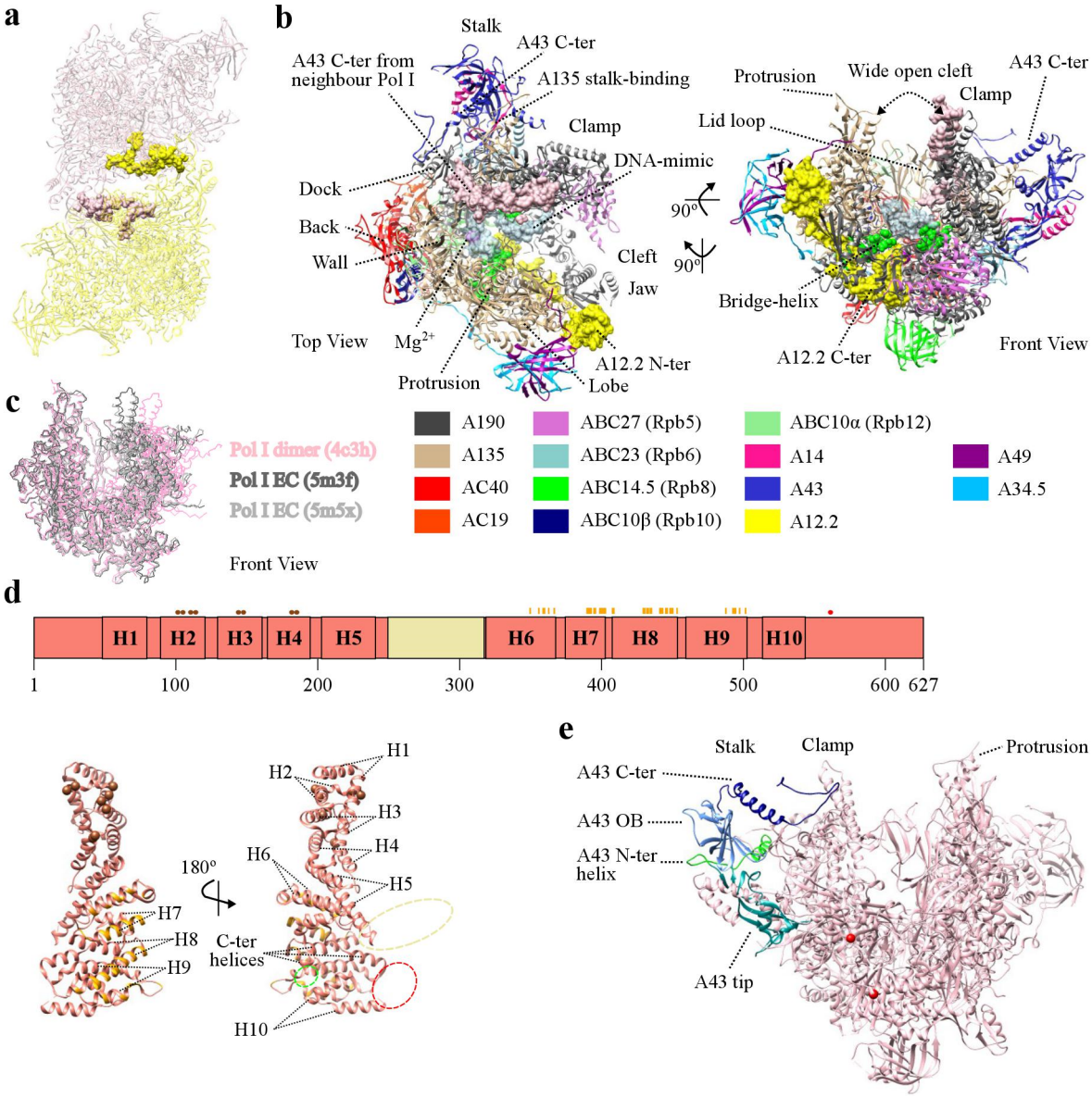

Figure 4. Structures of Pol I and Rrn3.

(a) X-ray crystal structure (pdb 4C3H) of inactive Pol I dimers. Monomers in yellow and pink with A43 C-ter in surface representation. (b) Dimeric Pol I (4C3H) crystal structure showing a single Pol I molecule. Surface representations: A43 C-ter tail of the neighbouring Pol I, pink; DNA-mimicking loop, light blue; bridge-helix partially unfolded, green; A12.2, yellow. Cleft aperture is incompatible with transcription. (c) Cleft closure in Pol I ECs. Pol I structures of inactive dimers and two ECs (Neyer et al 2016, Tafur et al 2016) superposed onto A135 (only A135 and A190 showed for clarity). (d) Diagram representation of Rrn3 and crystal structure of dimeric Rrn3 showing a single molecule (3TJ1, Blattner et al 2011). HEAT repeats (H1 to H10) are indicated. Pale yellow box and dashed ellipse, acidic loop; brown spheres, serine patch; red sphere, K558 crosslinked with Pol I A190 and AC40 subunits (Blattner et al 2011); orange rectangles and ribbon, dimerization interface; dashed red ellipse, unsolved 21-amino acid loop containing crosslinked K558; dashed green ellipse, Rrn3 C-ter. (e) A43 domains and lysines in Pol I (4C3H) crosslinked with Rrn3. Red spheres, lysines crosslinked with Rrn3 (K582 in A190 and K329 in AC40); green ribbon, A43 N-ter helix; dark cyan ribbon, A43 tip; light blue ribbon, A43 OB; dark blue ribbon, A43 C-ter.

bridge helix, which is a crucial element for catalysis and substrate incorporation in multi-subunit RNAPs (Weinzierl 2011) and appears unwound at its middle region (Figure 4b). This helix is totally folded in Pol I transcribing complexes (Neyer et al 2016, Tafur et al 2016). Presumably concomitant with partial bridge helix unfolding, a wide open cleft is found in the dimers (Figures 4b and 4c). Most interactions of A135 with nucleic acids in the vicinity of the active site, critical for NTP selection and translocation, and with the DNA:RNA duplex, cannot be established in a RNAP displaying this feature. In addition, and related to the previous inactivation characteristics of Pol I, the RNA exit channel is partially occluded by A190 lid loop.

Pol I atomic structures detailed other important Pol I feature, the location of A12.2 (Figure 4b). A12.2 comprises two Zn-binding β-ribbon domains connected by a flexible linker. The N-ter ribbon contacts the A190 jaw, the A135 lobe and the A49/A34.5 dimerization module in an equivalent location to Rpb9 N-ter in Pol II. However, the C-ter position of these subunits strongly differs. A12.2 C-ter Zn ribbon is located inside the NTP entry pore and places an acidic loop next to the active site, further reinforcing that A12.2 C-ter is involved in the second enzymatic activity of Pol I, its intrinsic ability to cleave RNA (Prescott et al 2004, Kuhn et al 2007). The residues correspondent to A12.2 acidic loop in TFIIS participate in RNA cleavage (Jeon et al 1994) and TFIIS and A12.2 C-termini are very similarly placed in Pol II and Pol I, respectively (Kettenberger et al 2003).

X-ray crystal structure of yeast Rrn3 at 2.8 Å resolution was solved (Blattner et al 2011) (Figure 4d). Rrn3 formed elongated dimers in the crystals and in solution, although binding Pol I as a monomer (Blattner et al 2011). The presence of Rrn3 dimers *in vivo* was not yet investigated. The structure revealed 10 HEAT* repeats formed by 20 anti-parallel α helices arranged in a superhelical fold. An acidic loop of approximately 70 amino acids between HEAT repeats 5 and 6, 47 N-ter and 10 C-ter residues were not solved but their deletions did not result in a growth phenotype *in vivo* (Blattner et al 2011). A 21-amino acid loop after HEAT repeat 10 completes the regions not observed in the crystal structure. Eight serine residues organised in 4 pairs (S101/S102, S109/S110, S145/S146 and S185/S186) stand on the N-ter surface of Rrn3, forming a serine patch partially conserved in higher eukaryotes. The C-ter half of the initiation factor contains residues composing a surface involved in homodimerization.

A structural model of Pol I:Rrn3 complex (Blattner et al 2011) was built taking into account three pieces of information. First, the interaction of Pol I A43 subunit with Rrn3 and their co-localization within Pol I:Rrn3 obtained from low resolution electron microscopy data (Peyroche et al 2000). Second, the effect of S145D and S185D mutations in Rrn3 serine patch, which impaired or strongly reduced Pol I:Rrn3 formation *in vitro*, respectively, and indicated the patch surface is most likely facing Pol I (Blattner et al 2011). Third, chemical crosslinking between lysine residues followed by mass spectrometry analysis identified probable interactions between Rrn3 around K558, contained

in the 21-amino acid loop, and regions in the vicinity of K582 in A190 and K329 in AC40 (Figures 4d and 4e). However, structural information on Pol I:Rrn3 complex detailing the interaction and not based on predicted models was lacking at the time this work was started.

6. The Recent Revolution in Single-Particle Cryo-EM

Cryo-EM and X-ray crystallography are commonly chosen techniques when it comes to determine the structure of large molecular assemblies. X-ray crystallography has been extensively used but, many times, high quality diffracting crystals are not obtained. Cryo-EM can circumvent this problem as it does not require crystallization. Cryo-EM is also free of staining agent and thus the resolution is not limited by the grain size like it happens in negative-staining EM. Cryo-EM rests in the pivotal discovery that very thin films of biological samples can be flash frozen in liquid ethane and stored embedded in amorphous ice (not crystalline ice) at liquid nitrogen temperatures (Adrian et al 1984), allowing the study of protein assemblies in a quasi-native state. Furthermore, two-dimensional (2D) electron micrographs must contain "single-particles" of individual biological assemblies in a range of random orientations enough to derive the three-dimensional (3D) structure.

Electron microscopes improved over the years (Henderson 2015): accelerating voltage has increased from 120 kV to 300 kV; thermionic emission electron guns were replaced by field emission guns, much brighter, therefore increasing electron beam coherence; temperature of the electron source cathodes dropped from approximately 3000°C to approximately 1800°C, increasing electron beam temporal stability; vacuum around cold specimens improved, reducing ice contamination and allowing data collection for days rather than hours (much more images recorded); stages stability greatly improved so that drift rates reduced. Furthermore, electron optics evolved to low hysteresis (multi lens condenser systems, for instance) (Schröder 2015) and automatic image acquisition made the data collection less laborious. However, since 2012 (Henderson 2015), two advances catapulted single-particle cryo-EM into a "resolution revolution" (Kühlbrandt 2014, Henderson 2015).

First, the development of direct detectors. These devices can integrate electrons without any additional conversion step (they are direct). CCD cameras need to convert the incident electron into photons, introducing additional noise, and films present higher detection times. Ergo, direct detectors can integrate the signal almost continuously (dozens of frames per second) without need for beam-blanking between frames. In practice, a "movie" is recorded and the individual frames stored. Such readout velocity could not be achieved before with CCD cameras or films. Also, direct detectors are much thinner than CCD, therefore reducing back-scattered electrons into neighbouring pixels (Schröder 2015). Direct detectors have a detective quantum efficiency (DQE) higher at 300 keV than film or CCD cameras (McMullan et al 2014) (higher DQE, higher signal-to-

noise ratio for the same electron dose exposure). In summary, direct detectors are superior to CCD cameras and film because they are faster and retrieve higher quality images. The possibility of storing and processing individual frames is associated to other important advance, the development of movie alignment software that compensates for beam induced movement (BIM) and restores high resolution detail in the images and in the final 3D reconstructions (Brilot et al 2012, Campbell et al 2012, Bai et al 2013, Li et al 2013, Allegretti et al 2014, Scheres 2014, Abrishami et al 2015, Rubinstein and Brubaker 2015).

The second major advance are the EM processing packages with maximum likelihood classification algorithms in two dimensions and, especially, in three dimensions (Scheres 2010, Sigworth et al 2010, Scheres 2012, Lyumkis et al 2013). These approaches stepped forward to deal with structural variability in the dataset. They successfully sort out 3D heterogeneity ranging from nonstoichiometric complex formation to large conformational changes or both. RELION (Scheres 2012), a recent program that quickly became widely used and incorporates maximum likelihood algorithms, generated higher resolution structures with old input data (des Georges et al 2014).

Advances in cryo-EM are being translated into a growing number of novel protein complex structures, sometimes at atomic resolution (Fernandez-Leiro and Scheres 2016). However, the full potential of single-particle cryo-EM was not yet achieved (Henderson 2015). It was estimated that 600 particles, independently of their size or molecular weight, would be enough to calculate a 3 Å structure (Rosenthal and Henderson 2003). Detectors DQE can still improve and BIM can be minimised once it is fully understood (Henderson 2015). New supporting materials as gold can contribute for BIM reduction (Russo and Passmore 2014). Also, microscopes equipped with energy filters and aberration correctors will provide better contrast images, perhaps pushing forward the resolution of 3D structures, and physical phase plates will enhance contrast and reduce the total number of particles to obtain a 3D reconstruction (Schröder 2015). Finally, computational replacement of central processing units by graphic processing units (GPUs) will accelerate the execution of the time-consuming maximum likelihood algorithms. The “resolution revolution” (Kühlbrandt 2014) is not concluded, it is happening.

OBJECTIVES

X-ray crystal structures of Pol I (Engel et al 2013, Fernández-Tornero et al 2013) corresponded to inactive dimers. The main objective of this work was the 3D structure determination at sub-nanometric resolution of free monomeric Pol I and Pol I:Rrn3 complexes to elucidate the structural changes in the transition from the inactive dimeric enzyme to the monomeric form and from monomers to activated Pol I. The enzyme responsible for 35S pre-rRNA production can adopt a set of different states prior to rDNA promoter recruitment and detailing them at the structural level deepens the knowledge on the mechanism of Pol I initiation.

These are the objectives:

- Study biochemical and functional properties of yeast Rrn3 *in vitro*, in particular if Rrn3 homodimers are disrupted by dilution and Rrn3 ability to bind rDNA.
- *In vitro* assembly of Pol I:Rrn3 complex.
- Determine the 3D structure of free monomeric Pol I, the most abundant Pol I population in the cell.
- Determine the 3D structure of Pol I:Rrn3, correspondent to the activated state of the enzyme.
- Obtain pseudo-atomic models for Pol I and Pol I:Rrn3 that enable comparisons between the different states of the enzyme: inactive dimers, monomers and activated Pol I bound to Rrn3.

MATERIALS AND METHODS

1. Materials

1.1. Media and Additives

Unless otherwise indicated, the media were prepared dissolving the components in water and autoclaving at 121°C for 20 min.

Table 2. Cell culture media.

Name	Composition
Lysogen broth (LB)	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl
LB plates	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar
Terrific Broth (TB) auto-inducible	1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 1x "TB salts"*, 1 mM $MgSO_4$*, 1x "5052"*, 1x NPS*
YPD	1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose**
YPDA	1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose**, 0.003% (w/v) adenine hemisulfate
YPD plates	1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose**, 2% (w/v) agar

*Stock solutions were 0.22 μm filter sterilised and added under sterile conditions.
**A 40% (w/v) glucose stock solution was prepared separately, autoclaved at 110°C and added under sterile conditions.

Table 3. Cell culture additives.

Stock	Work concentration
Ampicillin 1000x	100 μg/mL
Chloramphenicol 1000x	25 μg/mL
Kanamycin 1000x	50 μg/mL
$MgSO_4$ 1000x	1 mM
TB salts 10x	17 mM KH_2PO_4, 72 mM K_2HPO_4
5052 50x	0.5% (v/v) glycerol, 0.05% (w/v) glucose, 0.2% (w/v) α-lactose monohydrate
NPS 20x	25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4

1.2. Native Gels

Native gels were always run at 4°C in the same running buffer (25 mM Tris, 192 mM glycine, 2 mM DTT).

Table 4. Recipe for two native gels.
Gels with 0.75 mm spacer plates (BIORAD).

Stock	Percentage		
	4%	**6%**	**7%**
H_2O	6.4 mL	5.9 mL	5.7 mL
1.5 M Tris pH 8.8	2.5 mL	2.5 mL	2.5 mL
40% Acrylamide/Bis Solution 19:1 (BIORAD)	1.0 mL	1.5 mL	1.75 mL
APS 10 % (w/v) (prepared from APS powder, BIORAD)	100 µL	100 µL	100 µL
TEMED (BIORAD)	10 µL	10 µL	10 µL

2. Rrn3 Cloning, Expression and Purification

2.1. Rrn3 Cloning into pETM11

Rrn3 was amplified in 50 µL PCR reactions (Table 5) using an optimised temperature cycle (Table 6). 0.7 µL Pfu DNA polymerase 2.5 U/µL (Thermo Scientific) were added after the initial denaturation step (95°C, 5 min). The primers included NcoI (Forward - 5´cgcgccatggcggcttttgagaatacaagtaaacgaccac3`) and Acc65I (Reverse - 5´cgcgggtaccttattagtcatccgacccatcactttc3`) restriction sites.

Table 5. PCR mixture for Rrn3 amplification.

Yeast genomic DNA (10 ng/µL)	10 µL
dNTPs (10 mM each)	1 µL
Pfu buffer with $MgSO_4$ 10x	5 µL
Forward primer (10 µM)	5 µL
Reverse primer (10 µM)	5 µL
Sterile deionised H_2O	23.3 µL
Pfu DNA polymerase 2.5 U/µL	0.7 µL
(Total volume)	50 µL

Table 6. Temperature cycle used in Rrn3 PCR.

95°C	5 min		Initial Denaturation
(95°C	30 s		Denaturation
62°C	30 s		Primer annealing
72°C	5 min)	35 cycles	Extension
72°C	10 min		Final extension
10°C	∞		Storage

5 µL of the PCR reaction were run in an agarose gel to check amplification occurred. Four PCR reactions were loaded in a 0.8% agarose gel and the insert was purified from gel using High Pure PCR Product Purification Kit (Roche) and eluted in 50 µL Elution buffer (10 mM Tris-HCl pH 8.5 at 25°C). All purified DNA quantifications were done measuring the Abs_{260nm} in a spectrophotometer. Approximately 150 µg of gel purified Rrn3 were digested with NcoI/Acc65I restriction enzymes (New England Biolabs (NEB)) (Table 7) at 37°C overnight whereas NcoI/Acc65I digested pETM11 plasmid was available from the laboratory internal library of digested vectors. pETM11 is a non-commercial vector from the European Molecular Biology Laboratory (EMBL).

Table 7. Reaction mixture for Rrn3 digestion.

Purified Rrn3 PCR product	50 µL
Buffer 3 10 x (NEB)	15 µL
BSA 10x (NEB)	15 µL
NcoI 10 U/µL (NEB)	6.3 µL
Acc65I 10 U/µL (NEB)	6.3 µL
Sterile deionised H_2O	57 µL
(Total volume)	150 µL

After purification with High Pure PCR Product Purification Kit (Roche), the digested insert and vector were ligated overnight at 22°C using T4 DNA Ligase (Fermentas) and approximately a threefold molar excess of Rrn3. A control reaction without Rrn3 was set. Following standard molecular biology protocols, the whole ligation was transformed in homemade chemically competent DH5α cells which were plated in LB plates supplemented with kanamycin. The control reaction yielded no colonies. Some colonies from the ligation reaction were first digested with restriction enzymes to find positive digestion profiles and later sequenced (Secugen S. L.) to confirm mutations altering protein primary sequence were absent. Rrn3 cloning into pETM11 later resulted in the expression of Rrn3 N-terminally fused to a hexahistidine tag (His-tag) followed by a TEV protease recognition sequence.

2.2. Mutagenesis to Include a Strep-tag II® Fused to the C-terminus of Rrn3

A protocol based on QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies), although with modifications, was followed. First, a nicked circular plasmid containing the desired mutation was synthesized taking advantage of temperature cycling, a DNA polymerase and two complementary primers partially overlapping each chain of a template double-stranded plasmid. The non-overlapping region, located in the central part of each primer, contained the intended

insertion. Then, *Dpn I* enzyme digested the methylated template DNA, as it was isolated from an *E. coli* strain (DH5α) with Dam methylase, while leaving the synthesized mutant strands intact. Last, the nicked circular plasmid was transformed in competent cells with the ability to repair the nicks.

The template plasmid was the one used for Rrn3 expression (Rrn3 cloned into pETM11) and the coding sequence for a Strep-tag® II (5`WSHPQFEK3`, IBA) with a two residue linker (5′SAWSHPQFEK3′) was inserted after Rrn3 gene using primers 1 (5´gggaatatgaaagtgatgggtcggatgacagtgcttggagccacccgcagttcgaaaaataataaggtaccggatccgaattcgagc3`) and 2 (5´gctcgaattcggatccggtaccttattatttttcgaactgcgggtggctccaagcactgtcatccgacccatcactttcatattccc3`). The inserted mutation is underlined. A 50 µL reaction (Table 8) was subjected to an optimised temperature cycle (Table 9). 1 µL Pfu DNA polymerase 2.5 U/µL (Thermo Scientific) was added after the initial denaturation step (95°C, 1.5 min) (Table 9). N-ter His-tag and TEV protease recognition site in Rrn3 were kept and 5 µL of extension reaction were loaded in an agarose gel to check synthesis occurred.

Table 8. Reaction mixture for the synthesis of mutant plasmid strands.

Template plasmid (17.3 µg/mL)	0.5 µL
dNTPs (10 mM each)	1 µL
Pfu buffer with $MgSO_4$ 10x	5 µL
Primer 1 (2.5 µM)	1.5 µL
Primer 2 (2.5 µM)	1.5 µL
Sterile deionised H_2O	39.5 µL
Pfu DNA polymerase 2.5 U/µL	1 µL
(Total volume)	50 µL

Table 9. Temperature cycle used in the synthesis of mutant plasmid strands.

95°C	1.5 min		Initial denaturation
(95°C	50 s		Denaturation
58°C	60 s		Primer annealing
72°C	15 min)	18 cycles	Extension
72°C	10 min		Final extension
10°C	∞		Storage

1 µL Dpn I 20 Un/µL (New England Biolabs) was added to the remaining 45 µL extension reaction and the methylated DNA was digested 1 h at 37°C. 3 µL of the mixture were used to transform chemically competent DH5α cells, subsequently plated in LB plates supplemented with kanamycin.

Potential positive colonies were identified through restriction enzyme analysis and the absence of undesired mutations was confirmed by DNA sequencing (Secugen S. L.).

2.3. Rrn3 Expression

A protocol for Rrn3 expression was optimised based on the fact that yeast Rrn3 had been previously expressed at 24°C for 16 h in *E. coli* cells coding for rare codons and using an auto-inducible medium (Blattner et al 2011). Pre-inocula of BL21 (DE3) RIL (Stratagene) and Rosetta (DE3) (Novagen) cells transformed with the recombinant plasmid were grown overnight at 30°C in LB supplemented with kanamycin and chloramphenicol. In each case, 500 µL of the pre-inoculum were added to 500 mL TB auto-inducible and left at 37°C under agitation till the OD_{600nm} reached approximately 2, at which point the temperature was set to 24°C and the protein expressed for 16 to 18 h. The cultures were centrifuged 20 min at 6000 *g* and the cellular pellets stored at -80°C. Rrn3 soluble expression was assessed by sonication cell disruption followed by nickel resin pull down, SDS-PAGE analysis and coomassie staining. Both strains yielded similar amounts of soluble Rrn3. The protocol was scaled-up for 10 L using Rosetta (DE3) cells.

In view of Pol I:Rrn3 complex labelling, the described protocol was followed for the expression of 1.75 L of Rrn3 C-terminally fused to Strep-tag® II (Rrn3Strep2).

2.4. Rrn3 Purification

2.4.1. Rrn3 Employed in Pol I:Rrn3 Assembly

Protein was kept always on ice except for the chromatographic steps, run at RT with 0.22 µm filtered buffers previously cooled at 4°C. Cells correspondent to 5 L culture (45 g cell pellet) were thawed and suspended in 100 mL L buffer (50 mM HEPES pH 7.8, 200 mM NaCl, 10% (v/v) glycerol, 15 mM imidazol, 2 mM beta-mercaptoethanol) supplemented with protease inhibitors (cOmplete EDTA-free, Roche) and DNase (DNase I recombinant, RNase-free, Roche). Cells were sonicated (5 cycles with amplitude 40, process time 30 s, pulse-on time 10 s and pulse-off time 30 s) and the ice-cooled lysate was centrifuged at 20000 *g* for 40 min. The supernatant was loaded on a nickel affinity column (HisTrap, 5 mL, GE Healthcare) equilibrated in L buffer with 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and eluted in a linear gradient to 400 mM imidazole in 50 mL (10 column volumes (CV)). Rrn3-containing fractions, identified by SDS-PAGE analysis, were pooled, loaded on an anion-exchange chromatography column (Mono Q 5/50 GL, 1 mL bed volume, GE Healthcare) equilibrated in MQ Buffer (50 mM HEPES pH 7.8, 200 mM NaCl, 5 mM DTT) and eluted in a linear gradient to 750 mM NaCl in 20 mL (20 CV). After SDS-PAGE, pooled

fractions were concentrated and loaded on a gel filtration column (Superdex 200, 24 mL bed volume, GE Healthcare) equilibrated in GF buffer (20 mM HEPES pH 7.8, 100 mM Na_2SO4, 1 mM $MgCl_2$, 10 μM $ZnCl_2$, 5 mM DTT). Rrn3 fractions were concentrated to 20 mg/mL in a 30 kDa cut-off Amicon Ultra-15 centrifugal filter (Merck Millipore), frozen in liquid N_2 and stored at -80°C. Purified protein quantification was done measuring the Abs_{280nm} in a spectrophotometer.

2.4.2. Rrn3Strep2 Employed in the Assembly of Pol I:Rrn3:Antibody Complex

Isolation of Rrn3Strep2 followed essentially the same chromatographic steps, although some differences in the protocol were applied. The composition of the buffers was modified: cell pellet was suspended in L1 buffer (50 mM HEPES pH 7.8, 200 mM Na_2SO_4, 10% (v/v) glycerol, 15 mM imidazol, 2 mM beta-mercaptoethanol, 4 mM $MgCl_2$) supplemented with protease inhibitors (cOmplete EDTA-free from Roche, 0.5 mM PMSF) and DNase (DNase I recombinant, RNase-free, Roche); the nickel column was equilibrated in HT1 buffer (50 mM HEPES pH 7.8, 100 mM Na_2SO_4, 5% (v/v) glycerol, 15 mM imidazol, 2 mM beta-mercaptoethanol, 0.5 mM PMSF); the anion-exchange column was equilibrated in MQ1 buffer (25 mM HEPES pH 7.8, 100 mM Na_2SO_4, 5 mM DTT), same buffer used in the gel filtration. Also, the pellet of 1.75 L culture (approximately 13 g) was suspended in 50 mL L1 buffer, increasing the amount of buffer in relation to cells. Although Rrn3Strep2 was eluted from the nickel column equally applying a linear gradient to 400 mM imidazole, the gradient was done in 150 mL (30 CV). Finally, protein elution from the anion-exchange column was accomplished employing a linear gradient to 700 mM Na_2SO_4 in 30 mL (30 CV).

3. Sedimentation Velocity Assay

Rrn3 samples in U1 buffer (20 mM HEPES pH 7.8, 100 mM Na_2SO_4, 1 mM DTT) or U2 buffer (20 mM HEPES pH 7.8, 300 mM NaCl, 1 mM DTT) were loaded (320 μL) into analytical ultracentrifugation cells. The experiments were carried out at 43–48 krpm in a XL-I analytical ultracentrifuge (Beckman-Coulter Inc.) equipped with UV-VIS absorbance and Raleigh interference detection systems. Sedimentation profiles were recorded at 280 nm. Sedimentation coefficient distributions were calculated by least-squares boundary modelling of sedimentation velocity data using the continuous distribution c(s) Lamm equation model as implemented by SEDFIT 14.1 (Schuck 2000). Experimental *s* values were corrected to standard conditions (water, 4ºC, and infinite dilution) using the program SEDNTERP (Laue et al 1992) to get the corresponding standard *s* values ($s_{4,w}$).

4. Electrophoretic Mobility Shift Assay of Rrn3 and a Double-Stranded rDNA Fragment

Two reverse complementary DNA fragments ranging from position -50 to +50 in relation to the TSS (+1) were ordered (NTS fragment – 5´GGTTTAGTCATGGAGTACAAGTGTGAGGAAAAGTAGTTGGGAGGTACTTCATGCGAAAGCAG TTGAAGACAAGTTCGAAAAGAGTTTGGAAACGAATTCG3´; TS fragment – 5´CGAATTCGTTTCCAAACTCTTTTCGAACTTGTCTTCAACTGCTTTCGCATGAAGTACCTCCCAA CTACTTTTCCTCACACTTGTACTCCATGACTAAACC3´). 2 µM stock solutions were prepared in 0.22 µm filter-sterilised GF2 buffer (20 mM HEPES pH 7.8, 100 mM Na_2SO_4). Gloves were used whenever handling the DNA to avoid DNase contamination. The assembly of the rDNA fragment was accomplished by mixing NTS and TS chains in a 1:1 molar ratio (Table 10). The incubation and controls were heated in a water bath at 95°C for 3 min and gradually cooled down to RT. rDNA fragment formation was analysed in a 15% acrylamide gel.

Table 10. Reaction mixtures for rDNA fragment assembly and controls.

	rDNA fragment	NTS fragment control	TS fragment control
NTS fragment (2 µM)	5 µL (10 pmol)	5 µL (10 pmol)	----
TS fragment (2 µM)	5 µL (10 pmol)	----	5 µL (10 pmol)
GF2 buffer	----	5 µL	5 µL
(Final concentration (µM))	1 µM	1 µM	1 µM
(Final volume (µL))	10 µL	10 µL	10 µL

Rrn3:rDNA fragment incubation was done in a 1:1 molar ratio with appropriate controls (Table 11) and left reacting overnight at 4°C. Rrn3 N-terminally extended by a His-tag and a TEV protease recognition site presents a molecular weight of 75 kDa. The final buffer composition in Rrn3:rDNA incubation was 20 mM HEPES pH 7.8, 100 mM Na_2SO_4 with low concentrations of $MgCl_2$ (0.3 mM), $ZnCl_2$ (2.8 µM) and DTT (1.4 mM), present in GF buffer from Rrn3. Interaction was analysed in a 6% native gel detected with ethidium bromide and then silver-stained.

Table 11. Reaction mixtures for Rrn3:rDNA fragment and controls.

	NTS	Rrn3:NTS	TS	Rrn3:TS	Rrn3:rDNA	rDNA	Rrn3
Rrn3 0.2 mg/mL (2.7 µM)	----	1.9 µL (5 pmol)	----	1.9 µL (5 pmol)	1.9 µL (5 pmol)	----	1.9 µL (5 pmol)
rDNA fragment (1 µM)	----	----	----	----	5 µL (5 pmol)	5 µL (5 pmol)	----
NTS fragment (1 µM)	5 µL (5 pmol)	5 µL (5 pmol)	----	----	----	----	----
TS fragment (1 µM)	----	----	5 µL (5 pmol)	5 µL (5 pmol)	----	----	----
GF2 buffer	1.9 µL	----	1.9 µL	----	----	1.9 µL	5 µL
(Final volume)	6.9 µL	6.9 µL	6.9 µL	6.9 µL	6.9 µL	6.9 µL	6.9 µL

5. Pol I Endogenous Expression and Purification

5.1. Pol I Endogenous Expression

A *S. cerevisiae* strain (Euroscarf accession number SC1548, genotype *SC0000, MATa, ura3-52 leu2-3,112 YOR341w::TAP-KlURA3*) genetically modified to include a tandem affinity purification tag (TAP-tag) (Rigaut et al 1999) fused to the C-ter of A190 subunit was used in the endogenous expression of Pol I, intended to set crystallization trials and labelling of Pol I:Rrn3. Starting from glycerol stocks, fresh YPD plates were seeded and left 3 to 5 days at 30°C. Two big and isolated colonies were transferred to 5 mL YPDA and grown 24 h at 30°C and 200 rpm. Ampicillin was sometimes added to avoid bacterial contamination. The 5 mL were used to start a 500 mL YPDA culture and the cells were left 24 h at 30°C and 200 rpm. In the case of Pol I later used in crystallization trials, this last culture served to inoculate a 30 L fermenter whereas Pol I intended to Pol I:Rrn3 labelling was expressed inoculating 300 L of YPDA. Cells were pelleted in the transition from exponential to stationary phase, stored in dry ice for transportation and then at -80°C.

Pol I used in the EM reconstructions of Pol I:Rrn3 (not the labelled complexes) was expressed using a genetically modified strain to express a TAP-tag fused to the C-ter of AC40 subunit in Pol I and according to a published protocol (Moreno-Morcillo et al 2014).

5.2. Pol I Purification

5.2.1. Pol I Employed in Crystallization Trials

All chromatographic separations were run at 4°C and the 0.22 μm filtered buffers and protein fractions kept on ice throughout the purification. Approximately 600 g of cells (from two 30 L fermentations) were suspended in 650 mL L2 buffer (250 mM Tris-HCl pH 7.4, 30% (v/v) glycerol, 250 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 10 mM beta-mercaptoethanol, 50 mM sodium fluoride, 2 mM sodium pyrophosphate, 5 mM beta-glycerophosphate) supplemented with DNase (DNase I recombinant, RNase-free, Roche) and protease inhibitors (1 mM EDTA, 1 mM PMSF, 2 μg/mL leupeptin, 4 mM benzamidine, 1 μg/mL pepstatin). Lysis was accomplished in two passes in a GEA Niro Soavi homogenizer operated at 20000 psi and placed inside a 4°C chamber. The lysate was centrifuged for 1 h at 4°C, 30000 *g*, and the resulting supernatant loaded onto a manually assembled Heparin-Sepharose® Cl-6B column (cat no 17-0467, GE Healthcare, approximately 250 mL bed volume) equilibrated in Hep buffer (L2 buffer with 50 mM Tris-HCl pH 7.4, 5 mM beta-mercaptoethanol, 0.5 mM PMSF, 3 mM benzamidine and not supplemented with DNase). The column was washed first in Hep buffer and then in HepA buffer (Hep buffer with 2 mM beta-mercaptoethanol, 1 μg/mL leupeptin, 1 mM benzamidine, 0.7 μg/mL pepstatin and without glycerol). Protein was eluted using a linear gradient from 0% to 100% HepB (HepA buffer with 1 M $(NH_4)_2SO_4$). Heparin elution was incubated with 4 mL IgG Sepharose 6 Fast Flow resin (GE Healthcare) equilibrated in HepA, overnight at 4°C, in a Stuart® rotator (model SB3) set to 4 rpm.

The incubation was loaded onto a glass column (Glass Econo-Column, BIORAD), the resin washed with IgG buffer (50 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM $MgCl_2$, 10 μM $ZnCl_2$, 5 mM DTT, 50 mM sodium fluoride, 2 mM sodium pyrophosphate, 5 mM beta-glycerophosphate) and resuspended in approximately 15 mL IgG buffer. Taking advantage of the TEV protease recognition site in the TAP-tag, 300 μL homemade TEV protease (0.5 mg/mL) were added and the mixture incubated overnight at 4°C in a Stuart® rotator (model SB3) at 4 rpm.

Pol I in solution was recovered in 20 mL IgG buffer using the glass column and purified by anion exchange (Mono Q 5/50 GL, GE Healthcare). The column was equilibrated in MQ2 (40 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM $MgCl_2$, 10 μM $ZnCl_2$, 5 mM DTT) and the protein eluted using a gradient from 0% to 40% MQ3 (MQ2 with 1 M NaCl) in 20 CV. Protein fractions composition and quality were assessed by coomassie-stained 15% SDS-PAGE gels. Pools were formed and concentrated in 30 kDa cut-off Amicon Ultra-0.5 mL centrifugal filters (Merck Millipore). Aliquots were frozen in liquid N_2 and stored at -80°C. The purification protocol yielded 12-subunit Pol I deprived of the A49/A34.5 heterodimer (Pol I ΔA49/A34.5) and Pol I ΔA49/A34.5 further lacking the last 49 residues of A43 subunit (from here on named Pol I A43ΔCt ΔA49/A34.5). The protein

was stored at the NaCl concentration correspondent to anion-exchange chromatography elution, approximately 325 mM for Pol I ΔA49/A34.5 and 297 mM for Pol I A43ΔCt ΔA49/A34.5.

5.2.2. Pol I Employed in Pol I:Rrn3 Electron Microscopy Reconstructions

Pol I purification followed a described protocol (Moreno-Morcillo et al 2014) exploring the isolation properties of heparin, IgG and an anion-exchange column.

5.2.3. Pol I Employed in Pol I:Rrn3 Labelling

Compared to Pol I employed in crystallization trials and Pol I:Rrn3 EM reconstructions, the enzyme destined to the labelling experiment was isolated applying less chromatographic steps and using buffers with different composition. All chromatographic separations were run at 4°C and the 0.22 μm filtered buffers and protein fractions kept on ice throughout the purification. 500 g of pellet were suspended in 500 mL L3 buffer (200 mM HEPES pH 7.8, 25% (v/v) glycerol, 250 mM Na_2SO_4, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 10 mM beta-mercaptoethanol, 50 mM sodium fluoride, 2 mM sodium pyrophosphate, 5 mM beta-glycerophosphate) supplemented with DNase (DNase I recombinant, RNase-free, Roche) and protease inhibitors (cOmplete EDTA-free from Roche, 1 mM EDTA, 1 mM PMSF, 2 μg/mL leupeptin, 4 mM benzamidine, 1,4 μg/mL pepstatin). The cells were lysed in two passes at 20000 psi in a French® Press (Thermo Electron Corporation) equipped with a SLM-AMINCO pressure cell. The lysate was centrifuged 1 h at 25000 *g*, 4°C, in a J-LITE® JLA-16250 rotor (Beckman Coulter) and the supernatant carefully transferred into a cooled flask. Approximately 2.5 mL IgG Sepharose 6 Fast Flow resin (GE Healthcare), washed with deionized water and equilibrated in L3 buffer without DNase and protease inhibitors but EDTA, were incubated with the supernatant and left at 4°C, overnight, in a Stuart® roller mixer (model SRT9D) set to 12 rpm.

The whole total soluble protein incubation with IgG resin was centrifuged in two Falcon™ 50 mL Conical Centrifuge tubes at 4 ºC, 600 *g*, 1 min at a time, carefully pipetting the supernatant and centrifuging the incubation until all the resin deposited. Carefully re-suspending and centrifuging, the resin was washed three times with a total minimum of 20 CV of IgG1 buffer (50 mM HEPES pH 7.8, 10% (v/v) glycerol, 100 mM Na_2SO_4, 1 mM $MgCl_2$, 10 μM $ZnCl_2$, 5 mM DTT) and it was then resuspended in 15 to 20 mL IgG1 buffer. 800 μL TEV protease (0.5 mg/mL) were added and the mixture was left incubating overnight at 4 ºC and 4 rpm in a 50 mL Falcon™ sealed with Parafilm® and placed in a Stuart® rotator (model SB3).

Pol I released into solution was first recovered by centrifugation and then applying the supernatant into a cooled Poly-Prep® Chromatography Column (Bio-Rad). The filtered supernatant was

centrifuged 10 min at 14000 *g*, 4 ºC, in a JA-25.50 rotor (Beckman Coulter) and the operation repeated till a visible white precipitate, mainly TEV protease, was not detected. The protein was loaded in an anionic exchange column (Mono Q 5/50 GL, GE Healthcare) equilibrated in MQ4 buffer (20 mM HEPES pH 7.8, 5% (v/v) glycerol, 100 mM Na_2SO_4, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT), eluted running a gradient from 0% to 100% MQ5 (buffer MQ4 with 400 mM Na_2SO_4) at 0.5 mL/min in 3 mL (30 CV) and collected in 500 µL fractions. Fractions representing the complete chromatogram peak profile were run in a 15% SDS-PAGE and the gel was coomassie-stained to assess protein purity and help in the decision of the pools to form. A 4% native gel of selected fractions, run at 120 V, 4 ºC, for 2 to 3 h, loading 2 to 2.5 µg of protein and likewise coomassie-stained, was also useful in the decision of which fractions to join. The pools were formed, concentrated in a 30 kDa cut-off Amicon Ultra-0.5 mL centrifugal filter (Merck Millipore) and quantified before and after concentration to monitor protein recovery. Aliquots were made, frozen in liquid N_2 and stored at -80 ºC. The protein was stored at the salt concentration correspondent to anion-exchange chromatography elution. The purification protocol retrieved 14-subunit Pol I, 14-subunit Pol I lacking the last 49 residues of A43 (Pol I A43ΔCt), Pol I ΔA49/A34.5 and a minor population of Pol I A43ΔCt ΔA49/A34.5.

5.2.4. Western Blot of N-Terminally GFP-Tagged A43

Yeast whole cell lysates were run in a 12% SDS-PAGE gel. The protein was subjected to wet electrotransference (Mini-TransBlot Cell, BIORAD) at 30 V overnight using electrotransference buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol, 0.025% (w/v) SDS) and PVDF membrane (Sequi-BlotTM, BIORAD). The membrane was blocked 2 h in TBS-T (100 mM Tris, 137 mM NaCl, pH 7.6, 0.1% (v/v) Tween-20) 5% (w/v) milk at RT. After 5 washes of 5 min with TBS-T 0.5% (w/v) milk, the membrane was incubated with rabbit anti-GFP (Molecular Probes) and rabbit anti-Pgk1 (Invitrogen) antibodies in TBS-T 0.5% milk for 2 h at RT. Five washes of 5 min in TBS-T were done and GFP-tagged A43 and Pgk1 detected using a conjugated anti-rabbit IgG antibody and a chemiluminescent ECL Prime Western Blotting Detection Reagent (Amersham) following product instructions.

5.2.5. Cation-Exchange Chromatography to Resolve Pol I and Pol I ΔA49/A34.5

A mixture of 14-subunit Pol I and Pol I ΔA49/A34.5 (3.5 mg/mL, approximately 80 µg in total) was prepared in MS1 buffer (20 mM HEPES pH 7.8, 75 mM Na_2SO_4, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT). The mixture was injected in a cation-exchange column (Mono S PC 1.6/5, 0.1 mL bed volume, GE Healthcare) equilibrated in MS1 and the flowthrough was collected. A linear

gradient from 0% to 100% MS2 (MS1 with 500 mM Na_2SO_4) in 6 mL (60 CV) was applied and 50 µL fractions collected. Eluted protein was immediately analysed in coomassie-stained 15% SDS-PAGE gels. After overnight storage on ice at 4°C, pools of Pol I ΔA49/A34.5 and Pol I were formed, concentrated and analysed in 4% native gels stained with coomassie.

6. *In vitro* Assembly of Pol I:Rrn3 Complex

6.1. Assays to Demonstrate Pol I:Rrn3 Interaction

14-subunit Pol I or Pol I A43ΔCt were incubated overnight at 4°C with a molar excess of Rrn3 in I1 buffer (20 mM HEPES pH 7.8, 100 mM $(NH_4)_2SO_4$, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT). Pol I concentration in the incubation was 1 mg/mL. After centrifugation for 10 min at 10000 *g*, the mixture was run in a gel filtration column (Superdex 200 PC 3.2/30, GE Healthcare) and 50 µL fractions collected and analysed in a 15% SDS-PAGE. As controls, Pol I and Rrn3 were run separately in the amounts used in the incubation reaction.

The interaction was further confirmed by incubating 2 µg Pol I with Rrn3 in a 1:1 molar ratio overnight at 4°C in I1 or GF buffers. Pol I concentration ranged from 0.5 to 1 mg/mL. The Pol I:Rrn3 mixture and Pol I and Rrn3 controls were loaded in a 7% native gel subsequently coomassie-stained. Alternatively, 4% native gels were run and the protein electrotransferred (Mini-TransBlot Cell, BIORAD) overnight at 4°C and 30 V. The PVDF membrane (Sequi-Blot™, BIORAD) was blocked 2 h in TBS-T (100 mM Tris, 137 mM NaCl, pH 7.6, 0.1% (v/v) Tween-20) 5% (w/v) milk at RT. After 5 washes of 5 min with TBS-T 0.5% (w/v) milk, the membrane was incubated with anti-His antibody (cat no A7058, Sigma) (1:10000 dilution) in TBS-T 0.5% milk for 2 h at RT. 5 washes of 5 min in TBS-T were done and the His-tagged Rrn3 protein detected using chemiluminescent ECL Prime Western Blotting Detection Reagent (Amersham) following product instructions.

6.2. Assembly of Pol I:Rrn3 for Crystallization Trials

Both Pol I ΔA49/A34.5 and Pol I A43ΔCt ΔA49/A34.5 were used in crystallization trials. They were incubated with approximately a threefold molar excess of Rrn3 on ice overnight. Two incubation examples are detailed (Table 12). The molecular weight for Pol I ΔA49/A34.5 and Pol I A43ΔCt ΔA49/A34.5 was considered to be the same, 516.4 kDa, calculated after A49/A34.5 molecular weight subtraction from 590 kDa Pol I. Final Pol I concentration was 5 mg/mL and NaCl concentration lowered in relation to Pol I ΔA49/A34.5 and Pol I A43ΔCt ΔA49/A34.5 samples (approximately 325 mM and 297 mM, respectively, see section 5.2.1.). In Pol I ΔA49/A34.5:Rrn3

reaction, final buffer composition was 40 mM Tris pH 7.4, 193 mM NaCl, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT plus approximately 1.9 mM HEPES pH 7.8 and 9.6 mM Na_2SO_4. The last two components come from Rrn3 sample solution. Similarly, the final reaction buffer in Pol I A43ΔCt ΔA49/A34.5:Rrn3 was the same except for NaCl concentration (144 mM).

Table 12. Reaction mixtures for Pol I ΔA49/A34.5:Rrn3 and Pol I A43ΔCt ΔA49/A34.5:Rrn3.

	Pol I ΔA49/A34.5:Rrn3 (1:2.7 molar ratio)	Pol I A43ΔCt ΔA49/A34.5:Rrn3 (1:2.7 molar ratio)
Rrn3 20 mg/mL (270 µM)	6.1 µL (122 µg) (1.6 nmol)	7.3 µL (146 µg) (1.9 nmol)
Incubation1 buffer*	12.5 µL	13.6 µL
Incubation2 buffer**	14.9 µL	18.6 µL
Pol I ΔA49/A34.5 (10.6 mg/mL) (20.5 µM)	30 µL (318 µg) (0.6 nmol)	----
Pol I A43ΔCt ΔA49/A34.5 (10.3 mg/mL) (20.0 µM)	----	37 µL (381 µg) (0.7 nmol)
(Final volume)	63.5 µL	76.5 µL
(Final total Pol I concentration)	5 mg/mL	5 mg/mL
(Final NaCl concentration)	193 mM	144 mM

*Incubation1 buffer - 40 mM Tris pH 7.4, 200 mM NaCl, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT.
**Incubation2 buffer - 40 mM Tris pH 7.4, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT.

6.3. Assembly of Pol I:Rrn3 for Electron Microscopy

The sample was prepared similarly to Pol I:Rrn3:Antibody (see section 6.4), although with differences. Pol I was dialyzed against GF buffer before incubation with Rrn3 in a 1:1 molar ratio on ice overnight. The complex was crosslinked with 0.16% glutaraldehyde (final concentration) for 5 min on ice and subsequently quenched by addition of 1 M Tris-glycine pH 8.3. This incubation was injected in a Superdex 200 PC 3.2/30 (GE Healthcare) gel filtration column equilibrated in GF2 buffer (20 mM HEPES pH 7.8, 100 mM NaCl). The quality of the crosslinked sample was assessed by SDS-PAGE, native gel electrophoresis and liquid chromatography coupled to mass spectrometry.

6.4. Assembly of a Pol I:Rrn3:Antibody Complex for Electron Microscopy

For antibody labelling of Rrn3Strep2 in the context of Pol I:Rrn3 complex, Pol I and Rrn3Strep2 were mixed in a 1:1 molar ratio (Table 13) overnight at 4 ºC in 1.5 mL tubes sealed with Parafilm®. Pol I and Rrn3Strep2 stock dilutions, as well as the incubation dilution to a final Pol I

concentration of 2.5 mg/mL, were done with GF buffer (20 mM HEPES pH 7.8, 100 mM Na_2SO_4, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 5 mM DTT). The interaction occurred in GF buffer with glycerol traces and HEPES and Na_2SO_4 concentrations slightly above 20 mM and 100 mM, respectively, due to the contribution of original MQ1 buffer of Rrn3Strep2 and of MQ4/MQ5 buffer mixture of Pol I. 0.8 µL of the incubation and controls (2 µg Pol I) were loaded in a 4% native gel run 2 to 2.5 h at 4 ºC, 120 V, and silver stained to assess Pol I:Rrn3Strep2 complex formation.

Table 13. Reaction mixtures for Pol I:Rrn3Strep2 and controls.

	Pol I:Rrn3Strep2 (1:1 molar ratio)	Pol I control	Rrn3 control
Pol I (4 mg/mL)	-----	0.5 µL (2 µg)	-----
Pol I (20 mg/mL) (33.9 µM)	4.5 µL (90 µg) (153 pmol)	-----	-----
Rrn3Strep2 (0.5 mg/mL)	-----	-----	0.5 µL (0.25 µg)
Rrn3Strep2 (5.6 mg/mL) (75 µM)	2.1 µL (11.5 µg) (153 pmol)	-----	-----
GF buffer	29.5 µL	0.3 µL	0.3 µL
(Final volume)	36.1 µL	0.8 µL	0.8 µL
(Final total Pol I concentration)	2.5 mg/mL	2.5 mg/mL	-----

In order to avoid pH acidification induced by the fixing agent, HEPES final concentration was raised by adding a 0.5 M HEPES pH 7.8 solution. The sample was then crosslinked by the addition of glutaraldehyde 2.7% (v/v), prepared from the dilution of aqueous 25% glutaraldehyde (cat no G5882, Sigma-Aldrich) with GF buffer, to a final 0.5% (v/v) concentration. The reaction was allowed to continue for 5 min on ice, protected from light with an aluminium foil, and quenched by adding a solution of primary amines (25 mM Tris, 200 mM glycine, final concentrations) (Table 14). The volume of the crosslinking reaction correspondent to 2 µg of Pol I was loaded in a 4% native gel run 2 to 2.5 h at 4 ºC, 120 V, and silver stained. Fixation was also assessed in an 8% SDS-PAGE gel.

Table 14. Reaction mixture for the crosslinking of Pol I:Rrn3Strep2.

	Pol I:Rrn3Strep2 (crosslinked)
Pol I:Rrn3Strep2 incubation	35.2 µL
HEPES 0.5 M pH 7.8	8 µL
Glutaraldehyde 2.7% (v/v)	10.7 µL
Primary amines solution 10x (250 mM Tris pH 8.3, 2 M glycine)	6.1 µL
(Total volume)	60 µL
(Final total Pol I concentration)	1.5 mg/mL
(Final HEPES concentration)	82 mM

A 1.5 molar excess of the high-affinity Strep-tag® II specific monoclonal antibody (StrepMAB-Immo, IBA) was added to the crosslinked Pol I:Rrn3Strep2 (Table 15) to form the Pol I:Rrn3:Antibody complex. StrepMAB-Immo, lyophilized from a NH_4HCO_3 solution containing sucrose, was resuspended in AB buffer (50 mM HEPES pH 7.8, 100 mM Na_2SO_4). The assumed molecular mass for StrepMAB-Immo was 150 kDa.

Table 15. Reaction mixture for Pol I:Rrn3:Antibody and controls.

	Pol I:Rrn3:Antibody (1:1:1.5 molar ratio)	Pol I:Rrn3 control (crosslinked)	Antibody control
Pol I:Rrn3Strep2 (crosslinked) (approximately 1.5 mg/mL)	28.5 µL (72.5 pmol)	28.5 µL	-----
StrepMAB-Immo (1 mg/mL)	16.3 µL (109 pmol)	-----	16.3 µL
GF buffer	-----	16.3 µL	28.5 µL
(Total volume)	44.8 µL	44.8 µL	44.8 µL

Pol I:Rrn3:Antibody incubation was loaded in a Superose 6 PC 3.2/30 (GE Healthcare) equilibrated in GF1 buffer (20 mM HEPES pH 7,8, 85 mM Na2SO4, 3 mM DTT) to isolate Pol I:Rrn3:Antibody complex from the free antibody and crosslinking aggregates. The two controls were also loaded in the gel filtration column and the analysis of the three chromatogram profiles was useful for the identification of Pol I:Rrn3:Antibody enriched fractions. Representative fractions were analysed in 4% native gels run 2 to 2.5 h at 4 ºC, 120 V, and silver stained. The fractions were divided in two halves, one frozen in liquid N_2 and stored at -80 ºC and other kept at 4 ºC.

7. Crystallization Trials

All crystallization trials were carried out with a Cartesian Honeybee robot (Digilab) and set using sitting drop vapour diffusion technique in 96-well plates, mixing 0.1 µL reservoir solution with 0.1

μL protein solution. First, commercial sparse-matrix screens were tested at 22°C: for Pol I ΔA49/A34.5:Rrn3, Index (Hampton), JBScreen Classic 1, 2, 3 and 7 (Jena Bioscience) and Proplex, MemStart and MemSys (Molecular Dimensions) in a total of 384 conditions; in the case of Pol I A43ΔCt ΔA49/A34.5:Rrn3, JBScreen Classic 8 and 9 (Jena Bioscience), Morpheus (Molecular Dimensions), Low Ionic, Cryo, Protein-protein complexes and Membrane complexes (Sigma), Wizard I, II and III (Emerald BioSystems) and CP-Custom-II (Axygen) in a total of 582 conditions. Grid screen optimization at 22°C was tested for three initial hits without improvements, as well as 4°C crystal growth of grid screen duplicates. Efforts to crystallize Pol I:Rrn3 were ceased because negative-staining EM, that was being tested in parallel, gave promising results.

8. Electron Microscopy Sample Preparation, Data Collection and Processing

8.1. Negative-Staining Sample Preparation and Data Collection

3 μL of crosslinked Pol I or crosslinked Pol I:Rrn3 at concentrations ranging 100 to 150 μg/mL, or 3 μL Pol I:Rrn3:Antibody fractions eluted from the gel filtration, were adsorbed on glow-discharged carbon-coated copper grids (CF-400-Cu, Electron Microscopy Sciences) for 2 min. Sample excess was removed using a filter paper and the grids were quickly washed twice (2 x 4 μL H_2O). After the second wash, they were carefully dried in a filter paper and stained with 4 μL of 2% uranyl formate for 2 min.

Observations were performed in a JEOL-1230 electron microscope operated at 100 kV and micrographs recorded under low-dose conditions (approximately 10 $e^-/Å^2$) using a 4k x 4k TemCam-F416 camera (TVIPS). Pol I:Rrn3:Antibody micrographs were collected at 2.84 Å/pix. Crosslinked Pol I and crosslinked Pol I:Rrn3 data collection was performed at 2.28 Å/pix.

8.2. Negative-Staining Data Processing

With respect to crosslinked Pol I and crosslinked Pol I:Rrn3, 63500 images of Pol I and 42400 images of Pol I:Rrn3 were picked and extracted using EMAN (Tang et al 2007) and binned to 5.68 Å/pix for 2D analysis. The contrast transfer function (CTF) was estimated using CTFFIND3 and corrected by flipping phases. Reference-free averages were obtained using EMAN, while 3D reconstructions were carried out using protocols implemented in Xmipp (Scheres et al 2008). The correctness of the final structures was supported by the high correlation between 2D projections of the models and reference-free averages.

Pol I:Rrn3:Antibody data processing was entirely done inside Scipion (de la Rosa-Trevín et al 2016) using protocols from Scipion, CTFFIND3 (Mindell and Grigorieff 2003), Xmipp3 (de la Rosa-Trevín et al 2013) and RELION (Scheres 2012). Micrographs CTF estimation was done with CTFFIND3 and 15638 particles were manually picked and extracted from selected micrographs using Xmipp3. RELION was used for two rounds of 2D classification until unambiguous 2D averages containing the antibody were identified. A subset of 10997 particles corresponding to "good" averages (not necessarily showing density for the antibody) was subjected to RELION 3D classification using negative-staining Pol I reconstruction filtered at 60 Å as a reference, three classes and regularization parameter T = 2. One class exhibited extra density in Rrn3 region and served as a reference for a second 3D classification in two classes. A 3D class containing 1355 particles clearly corresponded to Pol I:Rrn3:Antibody complex.

8.3. Sample Vitrification and Data Collection

For cryo-EM analysis, 4 μL of crosslinked Pol I:Rrn3 at 60 ng/uL were placed onto glow-discharged Quantifoil R2/2 grids and incubated in the chamber of a FEI Vitrobot at 4 ºC and 95% humidity for 30 s before blotting for 2 s at an offset of -2 mm. The grids were rapidly plunged in liquid ethane and stored in liquid N_2 inside appropriate holders.

Data were collected on a FEI Titan Krios electron microscope operating at 300 kV, using FEI automated single particle acquisition software (EPU) on a back-thinned FEI Falcon II detector at a calibrated magnification of 79096 (pixel size of 1.77 Å). Defocus values in the final dataset ranged from 1.9 to 4.2 μm. Videos were intercepted at a rate of 68 frames for 4 s exposures.

8.4. Cryo-EM Data Processing

1288 movies of the vitrified Pol I:Rrn3 complex were averaged using optical flow correction (Abrishami et al 2015) and their CTF was estimated using CTFFIND4 (Rohou and Grigorieff 2015). Approximately 230000 particles were automatically selected using RELION (Scheres 2012), also employed for subsequent data processing. Two rounds of reference-free 2D class averaging allowed removal of bad particles, yielding a stack of 190750 good quality images. A negative-staining reconstruction of free Pol I was low-pass filtered at 60 Å and used as starting model to sort the images using 3D classification. Particles were split into 6 classes with T = 4, an offset search range of 6 pixels, and offset search steps of 2 pixels. Only one class containing a total of 32175 images clearly showed density for both Pol I and Rrn3, while a second class with 90173 particles corresponded to monomeric Pol I. Both particle sets were subjected to 3D refinement including particle polishing and post-processing, which yielded maps for monomeric Pol I and Pol

I:Rrn3 with final resolutions of 5.6 and 7.7 Å, respectively, according to the gold-standard Fourier shell correlation (FSC = 0.143). Finally, both particle sets were added together and subjected to the same refinement procedure, producing a map with a resolution of 4.9 Å.

9. Structure Modelling

The available crystal structure of dimeric Pol I (pdb 4C3H) was fitted into the cryo-EM map of monomeric Pol I at 4.9 Å resolution using UCSF Chimera (Pettersen et al 2004). Regions that appeared disordered (outside the cryo-EM map), were deleted from the model using COOT (Emsley and Cowtan 2004). The resulting model was divided into 30 different rigid bodies as previously defined (Moreno-Morcillo et al 2014), subjected to rigid-body real-space refinement using PHENIX (Adams et al 2010) and finally corrected for chain breaks. The same procedure was used for the cryo-EM map of the Pol I:Rrn3 complex by addition of the crystal structure of Rrn3 (3TJ1), which was divided into 2 rigid bodies at the disordered acidic loop in the middle region of the structure.

10. Data Availability

Pol I:Rrn3 and Pol I monomer cryo-EM maps (7.7 Å and 5.6 Å, respectively) were deposited under accession numbers EMD-4086 and EMD-4087. The cryo-EM map of Pol I monomer at 4.9 Å resolution and its corresponding pseudo-atomic model were deposited under accession codes EMD-4088 and PDB-5LMX.

11. Mass Spectrometry

11.1. Identification of Proteins by MALDI-TOF-TOF Peptide Mass Fingerprinting

The EXQuest Spot Cutter (BIO-RAD) was used for imaging the gels and picking the selected bands. For digestion, gel pieces were first washed with 50 mM ammonium bicarbonate (Sigma-Aldrich) and then with acetonitrile (ACN) (Scharlau). Trypsin (Promega) at a final concentration of 12.5 µg/mL in 50 mM ammonium bicarbonate solution was added to the gel pieces for 8 h at 37°C. Finally, 100% ACN containing 0.5% trifluoroacetic acid (TFA) (Sigma-Aldrich) was added for peptide extraction. Tryptic eluted peptides were dried by speed-vacuum centrifugation and re-suspended in 6 µL of 30% ACN + 0.1% TFA. 1 µL of each peptide mixture was deposited onto an 800 µm AnchorChip (Bruker-Daltonics) and dried at RT. 1 µL of matrix solution (3 mg/mL α-

cyano-4-hydroxycinnamic acid) in 33% ACN 0.1% TFA was then deposited onto the digest and allowed to dry at RT.

Samples were analysed with an Autoflex III TOF/TOF mass spectrometer (Bruker-Daltonics). Typically, 1000 scans for peptide mass fingerprinting and 2000 scans for MS/MS were collected. Automated analysis of mass data was performed using FlexAnalysis software (Bruker-Daltonics). Internal calibration of MALDI-TOF mass spectra was performed using two trypsin autolysis ions with m/z 842.510 and m/z 2211.105; for MALDI-MS/MS, calibrations were performed with fragment ion spectra obtained for the proton adducts of a peptide mixture covering the m/z 800–3200 region. The typical error observed in mass accuracy for calibration was usually below 20 ppm. MALDI–MS and MS/MS data were combined through the BioTools 3.0 program (Bruker-Daltonics) to interrogate the NCBI non-redundant protein database (NCBInr 20120318 (17574240 sequences; 6033299959 residues)) using MASCOT software 2.3 (Matrix Science). Relevant search parameters were set as follows: enzyme, trypsin; fixed modifications, carbamidomethyl (C); oxidation (M); 1 missed cleavage allowed; peptide tolerance, 50 ppm; MS/MS tolerance, 0.5 Da. Protein scores greater than 75 were considered significant ($p<0.05$).

11.2. Molecular Mass Determination

Experiments were performed on an Autoflex III MALDI-TOF-TOF instrument (Bruker Daltonics) with a smartbeam laser. The spectra were acquired using a laser power just above the ionization threshold. Samples were analysed in the positive ion detection and delayed extraction linear mode. Typically, 1000 laser shots were summed into a single mass spectrum. External calibration was performed using the bovine albumin (Sigma), covering the range from 15000 to 70000 Da. The 2,5-Dihydroxy-acetophenone (2,5-DHAP) matrix solution was prepared by dissolving 7.6 mg (50 μmol) in 375 μL ethanol followed by the addition of 125 μL of 80 mM diammonium hydrogen citrate aqueous solution. For sample preparation, 2.0 μL of the sample were diluted with 2.0 μL of 2% TFA aqueous solution and 2.0 μL of matrix solution. A volume of 1.0 μL of this mixture was spotted on the 800 μm AnchorChip target (Bruker-Daltonics) and allowed to dry at RT.

11.3. Liquid Chromatography Coupled to Mass Spectrometry Analysis

11.3.1. Pol I

Bands from 4% native gels were excised and processed manually. The digestion protocol was based on (Shevchenko et al 1996) with minor variations. Gel pieces were rinsed with 25 mM

ammonium bicarbonate and ACN 50:50 and dried under a stream of nitrogen. After reducing with 10 mM DTT (37 ºC, 30 min), samples were carbamidomethylated with 50 mM of iodoacetamide at RT for 20 min. Modified porcine trypsin from bovine pancreas (sequencing grade, Sigma) was added (100 ng per piece of gel) in a final volume of 20 µL of urea 2 M, 25 mM ammonium bicarbonate solution. The digestion took place overnight at 37 ºC, it was stopped by adding 0.5% TFA and peptides extracted immediately, dried by speed-vacuum centrifugation, resuspended in 5 µL of initial electrospray ionisation (ESI) solvent solution (100% water + 0.5% formic acid) and stored at -20 ºC until needed.

Liquid chromatography coupled to ESI mass spectrometry (LC ESI-MS/MS) analysis was performed using an Ultimate 3000 nanoHPLC (Dionex) coupled to an AmaZon Speed ion-trap mass spectrometer (Bruker Daltonics). The analytical column used was a silica-based reversed phase column C18 PepMap 75 µm × 15 cm, 3 µm particle size and 100 Å pore size (Dionex). The trap column was a C18 PepMap (Dionex), 5 µm particle diameter, 100 Å pore size, switched on-line with the analytical column. The loading pump delivered a solution of 0.1% TFA in 98% water/2% acetonitrile (LabScan) at 30 µl/min. The nanopump provided a flow-rate of 0.3 µL/min and was operated under gradient elution conditions, using 0.1% formic acid (Fluka) in water as mobile phase A, and 0.1% formic acid in 80% ACN/20% water as mobile phase B. Gradient elution was performed according the following scheme: isocratic conditions of 96% A: 4% B for 5 min, a linear increase to 40% B in 65 min, a linear increase to 95% B in 1 min, isocratic conditions of 95% B for 5 min and return to initial conditions in 2 min. Injection volume was 5 µl. The LC system was coupled via a nanospray source to the mass spectrometer. Automatic data-dependent acquisition using dynamic exclusion allowed to obtain sequentially both full scan (m/z 350-1500) MS spectra followed by tandem MS collision-induced dissociation (CID) spectra of the 4 most abundant ions.

MS and MS/MS data were used to search against a UniprotKB database (version 2013.06.26) containing 7800 *S. cerevisiae* entries. Database searches were done using a licensed version of Mascot v.2.2.04 (Matrix Science) and search parameters were set as follows: carbamidomethyl cysteine as fixed modification and oxidized methionine as variable one. Peptide mass tolerance was set at 0.5 Da both in MS and MS/MS mode, and 2 missed cleavages were allowed. In most cases, an accuracy of ± 0.1-0.2 Da was found both for MS and MS/MS spectra.

11.3.2. Crosslinked Pol I and Pol I:Rrn3

An overall similar protocol was adopted, although two important modifications were incorporated. First, as the fixation reaction affects lysines, trypsin was replaced by chymotrypsin (Roche Biochemicals) for the in-gel digestion according to (Shevchenko et al 1996). In summary, gel slices

were washed with 50 mM ammonium bicarbonate and samples reduced with 10 mM DTT. Alkylation was carried out with 55 mM iodoacetamide at RT before adding chymotrypsin (100 ng, Roche Biochemicals). The digestion took place overnight and the peptides were extracted, dried by speed-vacuum centrifugation and stored at -20 ºC until needed.

Second, LC ESI-MSMS analysis was performed using an Eksigent 1D-nanoHPLC coupled to a better performance 5600TripleTOF QTOF mass spectrometer (ABSciex). The analytical column used was a silica-based reversed phase column Waters nanoACQUITY UPLC 75 µm × 15 cm, 1.7 µm particle size. The trap column was a chromXP, 3 µm particle diameter, 120 Å pore size, switched on-line with the analytical column. The loading pump delivered a solution of 0.1% TFA in 98% water / 2% ACN (LabScan) at 3 µl/min. The nanopump provided a flow-rate of 0.25 µL/min and was operated under the same gradient elution conditions. Gradient elution was performed according the following scheme: isocratic conditions of 96% A: 4% B for 5 min, a linear increase to 40% B in 25 min, a linear increase to 95% B in 1 min, isocratic conditions of 95% B for 4 min and return to initial conditions in 10 min. Injection volume was also 5 µL, as well as the nanospray coupling between LC and the mass spectrometer. Automatic data-dependent acquisition using dynamic exclusion allowed obtaining full scan (m/z 350-1250) MS spectra followed by tandem MS CID spectra of the 25 most abundant ions.

Besides, protein identification was slightly different because a UniprotKB database (version 2014.01.01) containing 7802 *S. cerevisiae* entries was used for the search, peptide mass tolerance was set at 40 ppm and 0.1 Da for MS and MS/MS spectra, respectively, and 6 missed cleavages for chymotrytic peptides were allowed.

12. Protein Sequencing

Pol I samples were run in a 10% SDS-PAGE gel and the protein migrating in the region of interest was subjected to wet electrotransference (Mini-TransBlot Cell, BIORAD) for 1 h at 15°C, 100 V. The PVDF membrane (Sequi-Blot^{TM}, BIORAD) was stained in 1% (m/v) Coomassie Brilliant Blue R-250 (BIORAD), 10% (v/v) methanol for 1 min, unstained in 50% (v/v) methanol, washed 4 times in water and air dried. The stained target bands were cut from the PVDF membrane and subjected to automated Edman degradation (Blombäck et al 1966, Edman and Begg 1967) using a Procise pulsed-liquid protein sequencer (model 494, Perkin Elmer, Applied Biosystems Division) with an on-line 140C amino acid phenylthiohydantoin analyser.

RESULTS

1. Expression and Purification of Rrn3

Soluble yeast Rrn3 was overexpressed in *E. coli*. After affinity and anion-exchange chromatographies, purification of the initiation factor was polished in a gel filtration (Figure 5a). The maximum of the Abs_{280nm} for Rrn3 could not be precisely measured due to detector saturation. The elution volume (V_e), taken from the curve reading Abs_{260nm}, was 10.8 mL, value that may be compatible with an elongated dimer, the oligomeric state proposed for Rrn3 in solution (Blattner et al 2011). Rrn3Strep2 was similarly expressed and purified, although starting from less cellular pellet. The Ve, identical in the curves recording Abs_{280nm} and Abs_{260nm}, was 11.6 mL (Figure 5a). At the end of the protocol, Rrn3 and Rrn3Strep2 were pure, as assessed by SDS-PAGE (Figure 5b). The identity of the Pol I activating factor was further confirmed by western blot anti-His-tag and MALDI-TOF-TOF peptide mass fingerprinting. The isolation protocols yielded approximately 2 mg of Rrn3 per L of culture and about five times less (0.4 mg/L of culture) in the case of Rrn3Strep2.

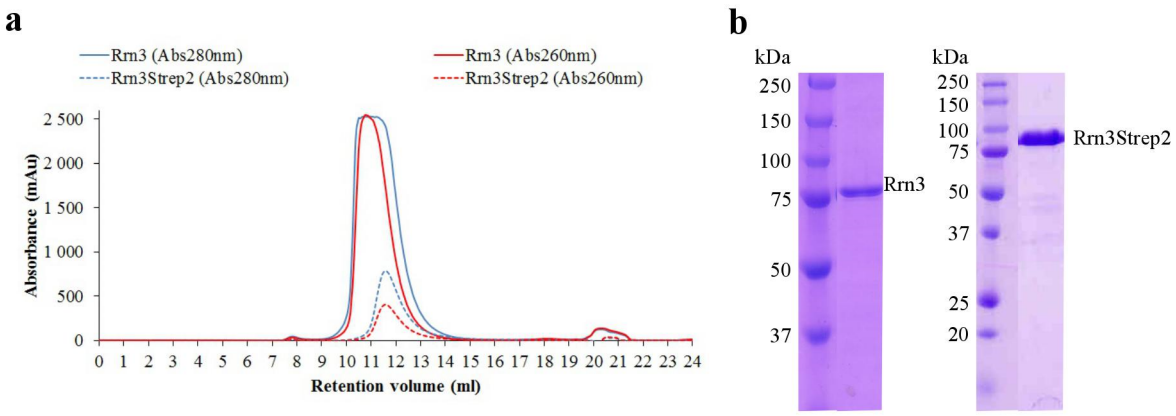

Figure 5. Purification of Rrn3 and Rrn3Strep2.
(a) Gel filtration of Rrn3 and Rrn3Strep2 (Superdex 200 10/300 GL). Elution volumes for Rrn3 and Rrn3Strep2 are 10.8 mL and 11.6 mL, respectively. (b) Coomassie-stained 15% SDS-PAGE gels of Rrn3 and Rrn3Strep2.

2. Rrn3 Exists in Solution Mainly as a Dimer and Dilution Favours a Monomer-Dimer Equilibrium

Using small-angle X-ray scattering (SAXS) technique, Rrn3 was found as a dimer in solution between 2 and 8 mg/mL (Blattner et al 2011). Sedimentation velocity assays were carried out to study Rrn3 oligomerization state in solution at different concentrations, ranging from 2 to 0.2 mg/mL. At 2 mg/mL and in 20 mM HEPES pH 7.8, 300 mM NaCl, 1 mM DTT, nearly 94% of the

total peak areas generated by Rrn3 corresponded to a peak of sedimentation coefficient (S) around 3.2, compatible with the theoretical molecular weight of an elongated Rrn3 dimer (Figure 6 and Table 16). At 1 mg/mL, nearly 82% of Rrn3 was found as a peak at 3.2 S and a new peak at 2.5 S, compatible with the mass of an Rrn3 monomer, corresponded to 11% of the total areas. At 0.2 mg/mL no defined peaks appear at 3.2 or 2.5 S, but rather a broad peak with a maximum at approximately 2.8 S, what might indicate that dimer-monomer equilibrium was established at this concentration. Sedimentation velocity results for Rrn3 at the same concentrations in 20 mM HEPES pH 7.8, 100 mM Na_2SO_4, 1 mM DTT gave similar results (Table 16). Thus, Rrn3 exists in solution mainly as a dimer at 2 mg/mL or higher concentrations and upon dilution dimer-monomer equilibrium might be gradually achieved.

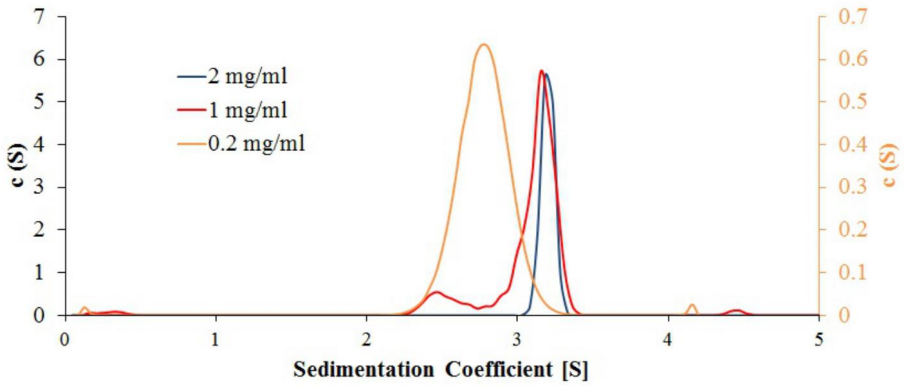

Figure 6. Rrn3 sedimentation velocity assay.
Rrn3 in 20 mM HEPES pH 7.8, 300 mM NaCl, 1 mM DTT. Orange curve represented in the secondary vertical axis.

Table 16. Summary of two Rrn3 sedimentation velocity assays.
Rrn3 in 20 mM HEPES pH 7.8, 300 mM NaCl, 1 mM DTT (abbreviated as NaCl) or 20 mM HEPES pH 7.8, 100 mM Na_2SO_4, 1 mM DTT (abbreviated as Na_2SO_4).

		Sedimentation coefficient (S)	
	Rrn3 concentration (mg/ml)	**2.5 S**	**3.2 S**
		Monomer	**Dimer**
NaCl	2.0	0%	94%
	1.0	11%	82%
	0.2	No defined peak	
Na_2SO_4	2.0	0%	98%
	1.0	8%	88%
	0.2	No defined peak	

3. Rrn3 Did Not Bind a rDNA Fragment in an Electrophoretic Mobility Shift Assay

Mammalian Rrn3 was shown to have a DNA-binding domain and independently interact with a 135 bps double-stranded rDNA fragment spanning TSS, property essential for rDNA transcription (Stepanchick et al 2013). In this work it was tested if a double-stranded yeast rDNA fragment of 100 bps (Figure 7a), ranging from position -50 to +50 in relation to the TSS, associated with Rrn3 *in vitro* to form an Rrn3:rDNA complex. After 1:1 molar ratio incubation overnight at 4°C in the presence of 100 mM Na_2SO_4, the incubation and controls were loaded in a 6% native gel and, first detected with ethidium bromide (nucleic acids sensitive), then silver stained (nucleic acids and protein sensitive) (Figure 7b).

Analysis of the silver staining reveals there is not a migration shift for Rrn3 in Rrn3:rDNA incubation, which migrates like the free Rrn3 control. Moreover, the signal for free rDNA fragment is identical in Rrn3:rDNA and in rDNA fragment control. With respect to ethidium bromide staining of nucleic acids, the detection profile for the rDNA fragment in the control and in the incubation is identical, reinforcing that the interaction does not occur. For the same reasons, Rrn3 does not bind the single-stranded chains. Thus, in an electrophoretic mobility shift assay (EMSA), yeast Rrn3 was not found to bind a 100 bps rDNA fragment spanning the TSS nor the single-stranded chains.

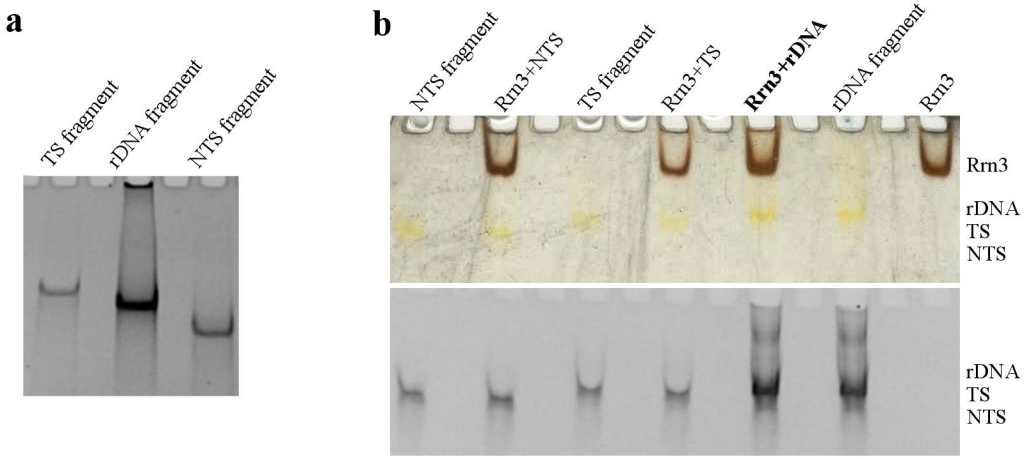

Figure 7. Electrophoretic mobility shift assay testing Rrn3 binding to a double-stranded rDNA fragment.
(a) 15% acrylamide native gel of double-stranded rDNA fragment assembly detected with ethidium bromide. TS, template strand; NTS, non-template strand. (b) 6% native gel of Rrn3 incubation with rDNA fragment and controls. Upper panel: silver staining. Lower panel: ethidium bromide detection.

4. Expression and Purification of Pol I

Pol I was endogenously expressed, as it would become too laborious to reconstitute the complex from the individual components or adopt co-expression strategies. As expected, the yield was significantly lower compared to overexpression protocols. Most successful purifications using A190-TAP strain yielded 42 to 50 μg of total Pol I per L of fermenter culture. First developed in yeast, the TAP-tag allows the isolation of protein complexes preserving their structural integrity (Rigaut et al 1999). Exploring the affinity of protein A in the TAP-tag for IgG, Pol I was purified in mild conditions. Last purification step was an anion-exchange chromatography, which allowed identification of different Pol I variants (Table 17). Importantly, an enriched 14-subunit Pol I population with all subunits intact, pure, was systematically obtained and constituted the most significant species isolated (Figure 8a, SDS-PAGE). Finally, inclusion of a cation-exchange chromatography as the final isolation step was assessed and revealed potential to resolve Pol I variants with distinct number of subunits.

Table 17. Pol I variants detected in the purification.
Pol I ΔA49 detected only after cation-exchange chromatography.

Nomenclature	Description
Pol I	14-subunit Pol I, complete
Pol I A43ΔCt	14-subunit Pol I lacking last 49 C-ter residues in A43 (A43Δ278-326)
Pol I ΔA49/A34.5	12-subunit Pol I lacking A49/A34.5 heterodimer
Pol I A43ΔCt ΔA49/A34.5	12-subunit Pol I lacking last 49 C-ter residues in A43 and A49/A34.5 heterodimer
Pol I ΔA49	13-subunit Pol I lacking A49

Observation of the anion-exchange chromatogram revealed two peaks (Figure 8a) and SDS-PAGE demonstrated the peak eluting at lower Na_2SO_4 concentration lacked full length A43 subunit. Furthermore, two bands seemed to appear in the A34.5 migrating region. Also, SDS-PAGE analysis of the corresponding anion-exchange peaks from a purification that yielded 12-subunit Pol I (Pol I ΔA49/A34.5) indicated A43 was migrating below AC40 in the peak eluted at lower NaCl concentration whereas it was full length in the major peak (Figure 8b).

At this stage, representative fractions of the two peaks (Figure 8a) containing Pol I in solution were sent to molecular weight analysis by MS-MALDI TOF. A molecule with m/z coincident with full length A43 (36.2 kDa) was not detected in any of the peaks whereas a molecule of 30.6 kDa was found only in the lower salt elution peak. The m/z of all other molecules in the studied range coincided for the two anion-exchange peaks. Since acidic proteins are many times difficult to detect in MALDI-MS (Rodthongkum et al 2010), a possible explanation for the failure to detect full length A43 in the higher salt elution peak could reside in the acidic C-ter region of this subunit.

On the other hand, the 30.6 kDa molecule exclusively found in the lower salt elution peak could correspond to A43 lacking 49 residues from the acidic C-ter.

Protein sequencing of the electrotransferred SDS-PAGE bands (full length A43 and A43 migrating near A34.5) showed that in both situations the first 4 N-ter residues were SQVK, being S the amino acid right after the initial methionine. Besides a likely removal of the first methionine of A43 in yeast, facilitated if the second amino acid has a small enough side chain as in the case of S (Frottin et al 2006, Bonissone et al 2013), this result demonstrated the N-ter of both proteins was identical and the difference between them did not lay in N-ter degradation. Together, the data on the anion-exchange chromatography, SDS-PAGE, molecular weight analysis by MS-MALDI TOF and protein sequencing constituted solid evidence that the lower salt elution peak corresponded to 14-subunit Pol I lacking 49 residues in A43 C-ter (Pol I A43ΔCt) while the higher salt elution peak contained 14-subunit Pol I with all subunits intact.

Based on the fact that Pol I A43ΔCt was systematically obtained, it was investigated if A43 C-ter was lost in the course of the isolation procedure or if there was A43ΔCt inside yeast cells. In collaboration with Olga Calvo`s laboratory, a yeast strain with GFP N-terminally fused to A43 was cultured and cells lysed at exponential and stationary growth in a buffer containing protein denaturing agents to avoid proteolysis by an hypothetical A43-degrading protease. Following SDS-PAGE and GFP and Pgk1 immunodetection (anti-Pgk1 as loading control), A43 levels in yeast whole cell lysates were analysed (Figure 8c). In exponential phase, the total amount of A43 is kept stable and A43ΔCt is only residually detected. In stationary phase, A43 reduces and A43ΔCt seems to increase accordingly. This result suggests Pol I A43ΔCt might exist in living yeast cells, especially in stationary phase.

When loading Pol I or Pol I A43ΔCt in a native gel, two predominant bands were observed (Figure 8d, left panel). To elucidate the nature of each Pol I population, the bands were separately analysed by LC ESI-MS/MS for protein identification. In the lower band all subunits but A49 and A34.5 were detected whereas in the upper band the heterodimer was identified, showing Pol I ΔA49/A34.5 and 14-subunit Pol I populations co-exist in both peaks eluted from the anion exchange column. It was previously shown that A49 and A34.5 can dissociate from Pol I (Huet et al 1976, Kuhn et al 2007) and native MS data identified complete Pol I and Pol I ΔA49/A34.5 in samples of the enzyme (Geiger et al 2010). Furthermore, A49/A34.5 heterodimer is the only sub-complex to dissociate from Pol I when decreasing the ionic strength from 800 mM to 160 mM NH_4CH_3COO (Geiger et al 2010). Also, for reasons not fully understood but most likely related to unintended and harsher conditions applied in the purification protocol, there were Pol I purifications yielding only 12-subunit Pol I (Pol I ΔA49/A34.5 and Pol I A43ΔCt ΔA49/A34.5 (Figure 8b)).

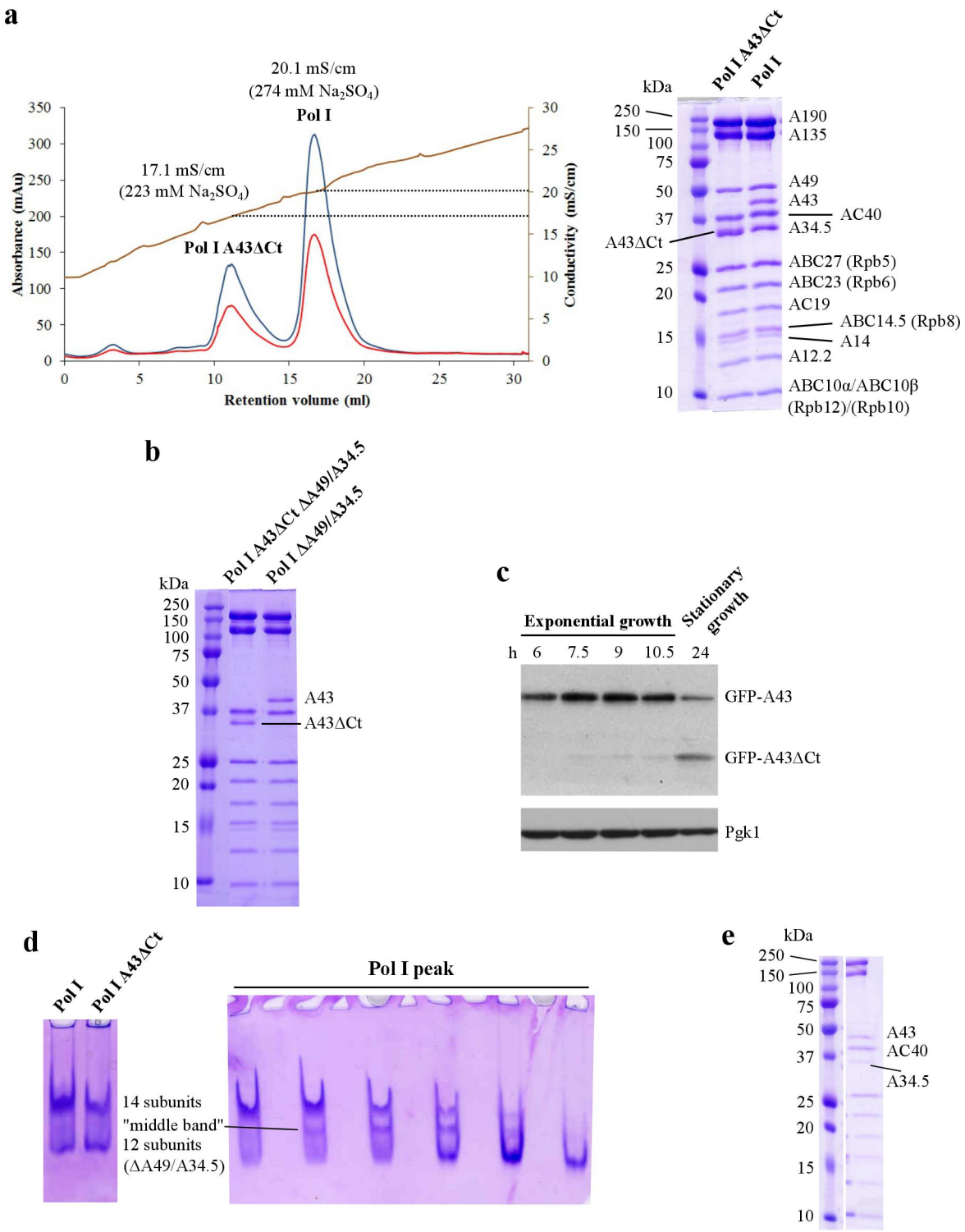

Figure 8. Pol I purification.

(a) Anion-exchange chromatography of Pol I and coomassie-stained 15% SDS-PAGE gel. In the chromatogram, Abs_{280nm} and Abs_{260nm} curves are in blue and red, respectively, and conductivity in brown. (b) Coomassie-stained 15% SDS-PAGE gel of Pol I ΔA49/A34.5 and Pol I A43ΔCt ΔA49/A34.5. (c) Yeast whole cell lysates from exponential and stationary growth phases analysed by western blot using antibodies against GFP (A43 and A43ΔCt detection) and Pgk1 (loading control). (d) Coomassie-stained 4% native gels of Pol I, Pol I A43ΔCt and fractions of Pol I peak in (a). (e) Coomassie-stained 15% SDS-PAGE gel of Pol I enriched in the native gel "middle band" and Pol I ΔA49/A34.5 band (poor in 14-subunit Pol I).

Interestingly, native gels loading individual fractions inside each peak showed the ratio between Pol I and Pol I ΔA49/A34.5 is higher in the ascending part of the curve and then gradually decreases till low content protein fractions in the descending curve present only Pol I ΔA49/A34.5 (Figure 8d, right panel). Thus, anion exchange chromatography, while efficient in resolving Pol I and Pol I A43ΔCt, did not completely resolve 14-subunit Pol I and Pol I ΔA49/A34.5. Additionally, in the native gels, an extra band between 14-subunit Pol I and Pol I ΔA49/A34.5 was often visible (Figure 8d, "middle band"). SDS-PAGE analysis of fractions enriched in this band (but also containing Pol I and Pol I ΔA49/A34.5) revealed A49 was practically absent (its amount not enough for coomassie detection) and A34.5 was also sub-stoichiometric (Figure 8e). Conceivably, the extra band in the native gel corresponds mainly to Pol I ΔA49 in co-migration with Pol I ΔA34.5. It is possible that A34.5 binds Pol I without A49 and vice-versa, as both subunits present extended arms fixing them to A135 in Pol I (Fernández-Tornero et al 2013).

A mixture of 14-subunit Pol I and Pol I ΔA49/A34.5 was loaded in a cation-exchange column and the fractions analysed by SDS-PAGE (Figure 9). The flowthrough contained Pol I lacking the heterodimer A49/A34.5. The peak eluted at lower Na_2SO_4 concentration corresponded to Pol I lacking A49. The higher ionic strength peak was not sharp and symmetrical, rather exhibiting a "shoulder" in the ascending part of the curve, suggesting mixture of protein complexes. SDS-PAGE indicated that right after elution A49 and A34.5 are present and seem stoichiometric. Most likely, the major population in the peak is intact 14-subunit Pol I. After keeping the protein overnight on ice, pool formation representing each peak and concentration, a native gel was run (Figure 9, native gel). While Pol I ΔA49/A34.5 remained stable, the 14-subunit enzyme exhibited two bands nearly the same intensity. This result might be explained by an effect of the native gel protocol or the concentration step on Pol I stability. A possible equilibrium between Pol I and Pol I ΔA49/A34.5, or even Pol I post-translational modifications, cannot be discarded, as well as different contributions of the hypothesis formulated. In summary, a cation-exchange column is able to resolve Pol I ΔA49/A34.5, Pol I ΔA49 and Pol I, although the 14-subunit Pol I peak is not sharp and presents a "shoulder". The Pol I ΔA49/A34.5 population was proved to be stable while further investigation would be required to understand 14-subunit Pol I behaviour over time.

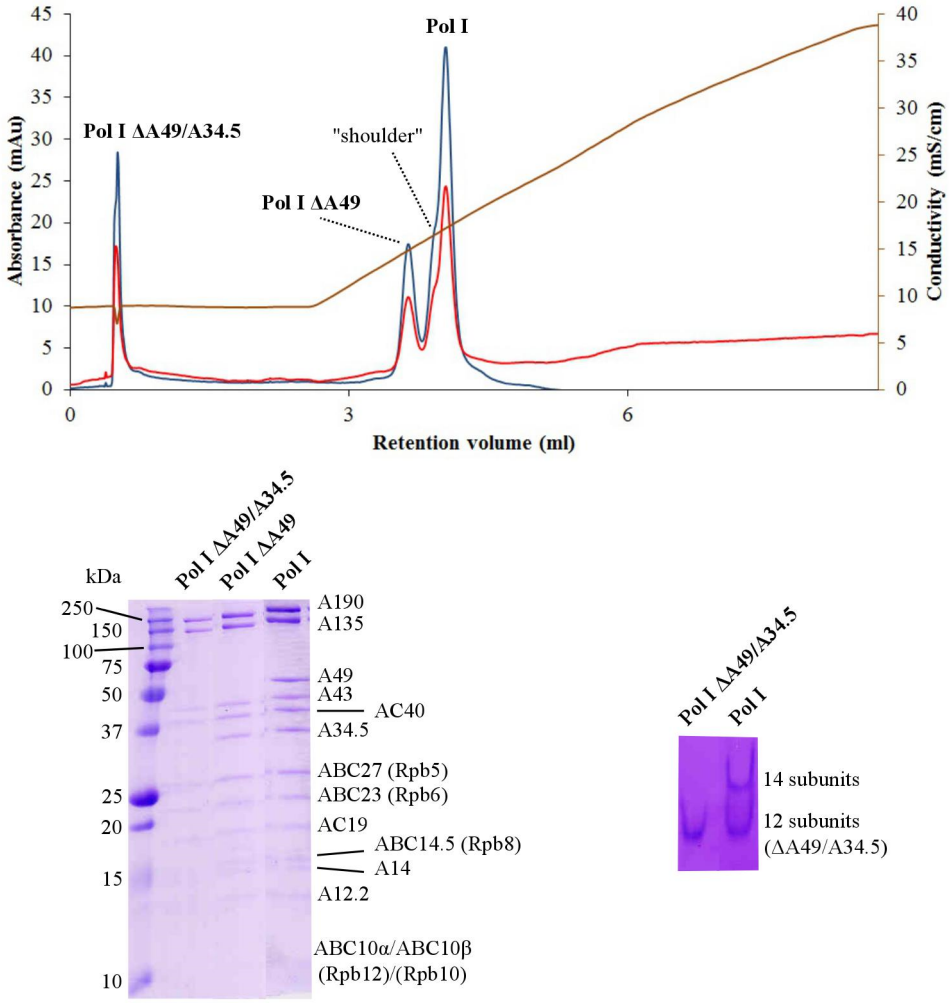

Figure 9. Cation-exchange chromatography of a mixture of 14-subunit Pol I and Pol I ΔA49/A34.5.
In the chromatogram, Abs_{280nm} and Abs_{260nm} curves are in blue and red, respectively, and conductivity in brown. Coomassie-stained 15% SDS-PAGE (bottom left) and 4% native (bottom right) gels.

5. *In vitro* Assembly of Pol I:Rrn3 Complex

In order to check Pol I:Rrn3 complex formation, Pol I was incubated overnight at 4°C with a molar excess of Rrn3, the mixture run in a gel filtration column (Superdex 200 PC 3.2/30, GE Healthcare) and the fractions analysed in a 15% SDS-PAGE gel (Figure 10a). As controls, Pol I and Rrn3 were run separately in the amounts used in the interaction reaction. Rrn3 and the 14 subunits of Pol I were detected in the peak with an elution volume (V_e) of 1.05 mL (Figure 10a, red curve). Furthermore, there was a slight shift in the V_e compared to the Pol I control (1.05 mL versus 1.07 mL, compare red and grey curves), suggesting a bigger complex than Pol I was formed in the incubation reaction. Also, the 1.05 mL V_e peak presented higher Abs_{280nm} than Pol I control peak (V_e = 1.07 mL). Finally, the peak for free Rrn3 (V_e = 1.32 mL) presented a higher maximum in the Rrn3 control (pink curve) than in the incubation chromatogram (red curve), suggesting Rrn3 was

displaced to interact with Pol I. Pol I:Rrn3 interaction was accomplished using Pol I purified in the presence of phosphatase inhibitors or in their absence.

Using the same protocol but greater molar excess of Rrn3, Pol I A43ΔCt was shown to interact with Rrn3 (Figure 10a). In conclusion, Pol I or Pol I A43ΔCt incubation with Rrn3 followed by gel filtration separation demonstrated Pol I:Rrn3 and Pol I A43ΔCt:Rrn3 could be assembled *in vitro* and that the C-ter (last 49 residues) of A43 was not required for activated complex formation. An important A43 function resides in the interaction with Rrn3 and the A43 C-ter is dispensable for that association.

The interaction was further demonstrated by the overnight incubation of Pol I and Rrn3 at 4°C in a 1:1 molar ratio followed by analysis in a 7% native gel. Rrn3 and Pol I in the amounts used in the incubation were loaded as controls (Figure 10a, inset). It can be observed that some free Rrn3 is still detected in the incubation, suggesting that Pol I:Rrn3 interaction is an incomplete reaction and at least three distinct populations are present: Pol I:Rrn3, Pol I and Rrn3. A similar conclusion can be drawn when detecting Rrn3 with an anti-His-tag antibody in a PVDF membrane blotted from a 4% native gel (data not shown). Rrn3 is detected most likely co-migrating with Pol I in Pol I:Rrn3 complex but also in its fast-migrating free form.

The incompleteness of Pol I:Rrn3 assembly reaction urged the use of a fixation technique in search for an enrichment of Pol I:Rrn3 complex. Distinct glutaraldehyde concentrations were first tested in free Pol I and the fixation assessed by SDS-PAGE and native gels (Figure 10b). It was chosen the lowest glutaraldehyde concentration (% (w/v)) able to prevent Pol I entry in the SDS-PAGE gel and, simultaneously, generating a band in the native PAGE compatible with a free monomeric state, without lysine crosslinking between distinct enzyme molecules (Figure 10b). Due to glutaraldehyde, fixed Pol I was interpreted to be a fast-migrating band.

Concentrations near the lowest found to be effective were later used, especially in Pol I:Rrn3 cryo-EM sample (0.16% glutaraldehyde). Thus, Pol I:Rrn3 complex was crosslinked with glutaraldehyde seeking preservation of the macromolecular protein assembly in a native state. Chemically fixed Pol I:Rrn3 was loaded in a 8% SDS-PAGE gel (Figure 10c). Individual Pol I subunits and Rrn3 did not enter the running gel and, in turn, protein was detected in the low-acrylamide percentage stacking gel, suggesting that chemical fixation occurred. Still, a 4% native gel (Figure 10c) was also run to confirm the existence of Pol I:Rrn3 without intermolecular crosslinking but rather intramolecularly fixed, moving below crosslinked aggregates. Similar to Pol I, crosslinked Pol I:Rrn3 migrated faster than Pol I:Rrn3 not fixed. LC ESI-MSMS confirmed the presence of Rrn3, A49 and A34.5, as well as of the remaining 12 subunits of Pol I.

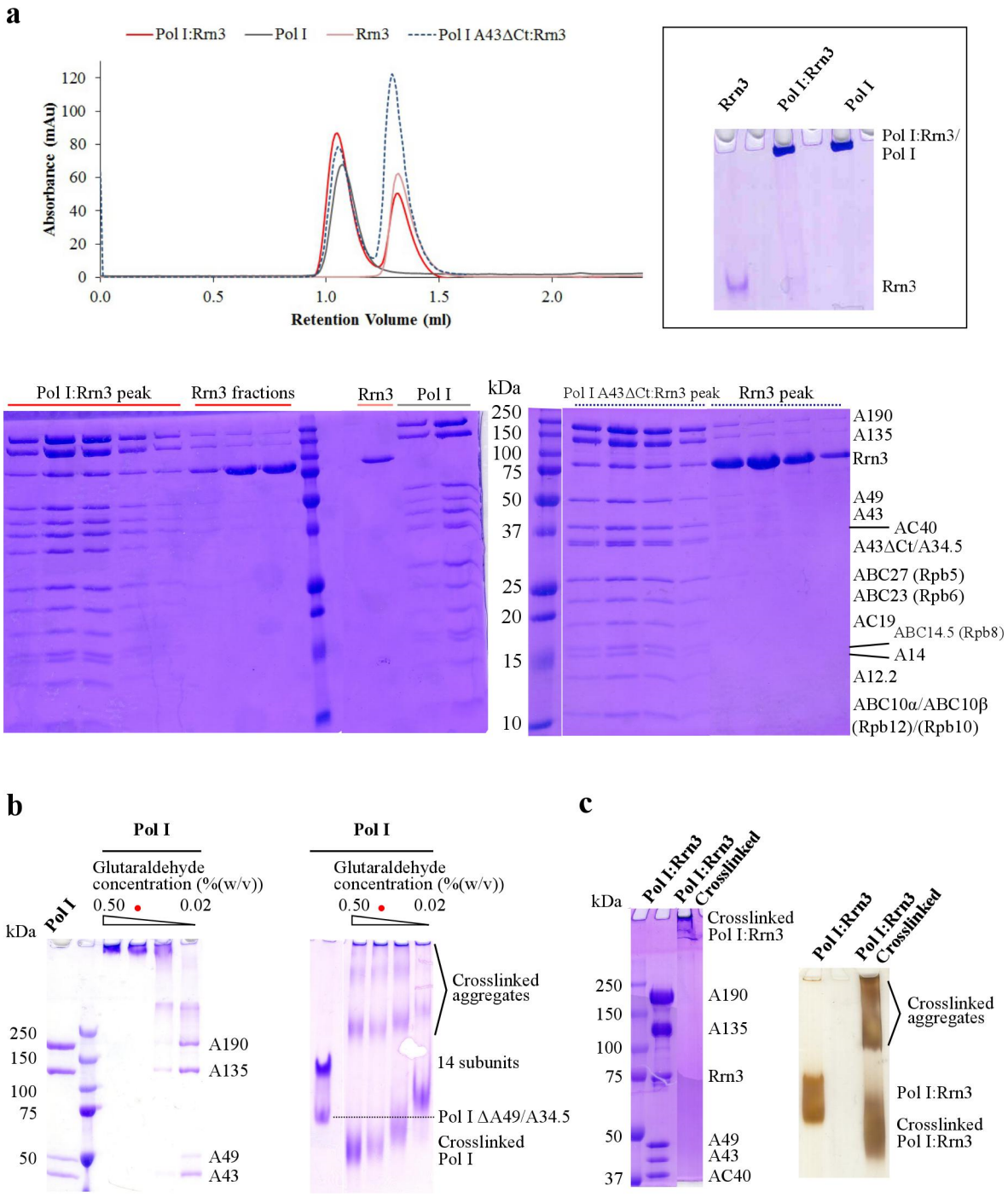

Figure 10. *In vitro* assembly of Pol I:Rrn3.

(a) Pol I:Rrn3 and Pol I A43ΔCt:Rrn3 interaction assessed by gel filtration and coomassie-stained 15% SDS-PAGE gels. Only the Abs_{280nm} is represented in the chromatograms. The inset in the top right corner displays Pol I:Rrn3 and controls in a coomassie-stained 7% native gel. (b) Titration of Pol I with glutaraldeheyde assessed in coomassie-stained 8% SDS-PAGE (left) or 4% native (right) gels. Red spheres signal the chosen glutaraldeheyde concentration. (c) Pol I:Rrn3 crosslink assessed in an 8% SDS-PAGE gel (left) and in a silver-stained 4% native gel (right).

6. *In vitro* Assembly of Pol I:Rrn3:Antibody Complex

To elucidate Rrn3 orientation in Pol I:Rrn3 reconstruction at intermediate resolution, different methods for Rrn3 labelling were considered. Preliminary results of histidine-tagged Rrn3 labelling with nickel (II) conjugates showed a reduced efficiency, as a small amount of total Rrn3 was labelled or the labelling was poorly detected. StrepMAB-Immo (IBA) is presented as a high-affinity Strep-tag® II specific monoclonal antibody especially suited for capturing of Strep-tag® II fusion proteins on solid phases, used, for instance, in surface plasmon resonance. The idea of using the antibody in EM protein labelling came, in part, from its high affinity, but also as an alternative to nickel (II) conjugates to fluorescent dyes or gold. An initial test of the association between Rrn3Strep2 and the antibody, assessed by EMSA, confirmed the high affinity of the interaction and prompted the labelling assays including Pol I.

To investigate the antibody binding to Rrn3 in the context of Pol I:Rrn3, Pol I was incubated with Rrn3Strep2 for 6 h at 4°C, then the antibody was added and left incubating overnight (1:1:1 molar ratio). The incubation and controls were loaded in a native gel (Figure 11a). As initially observed, Rrn3Strep2 interacted with the antibody because there was a pronounced migration shift if Rrn3 and Rrn3:Antibody wells were compared. The association seemed to be maintained in the triple incubation with formation of Pol I:Rrn3:Antibody, as at least an additional slow migrating band emerged in relation to Pol I:Rrn3, Pol I or Pol I:Antibody controls. The extra band attributed to the labelled Pol I:Rrn3 gave an acceptable signal in terms of intensity and justified the antibody choice as a labelling tool. Approximately 30% of the protein in the sample was estimated to be Pol I:Rrn3:Antibody.

Once attested the formation of Pol I:Rrn3:Antibody, the next step was crosslinking the ternary complex and study the sample by negative-staining EM. However, evidence for the presence of the antibody was never convincing enough in the 2D averages or 3D reconstruction. The second and successful approach was to reconstitute Pol I:Rrn3, fix this complex and only then add the antibody. In this case, to favour its interaction with Pol I:Rrn3, the antibody was used in a 1.5 molar excess in relation to Pol I and Rrn3. In order to remove crosslinking aggregates and the excess of antibody, the incubation was loaded in a gel filtration column (Superose 6 PC 3.2/30, GE Healthcare), as well as the crosslinked Pol I:Rrn3 control (Figure 11b). In each chromatogram, the peak with the highest absorbance maximum is found in the void volume and corresponds to crosslinking aggregates. Peaks with the second higher absorbance maximums present different elution volumes, with a shift towards a higher molecular weight in the Pol I:Rrn3:Antibody chromatogram. Analysis of a silver-stained native gel containing eluted fractions from the incubation and control (Figure 11b) allowed the identification of an extra band not present in crosslinked Pol I:Rrn3, attributed to crosslinked Pol I:Rrn3:Antibody. The fraction most enriched

in this band (identified in Figure 11b by a single blue sphere) was used in negative-staining EM studies.

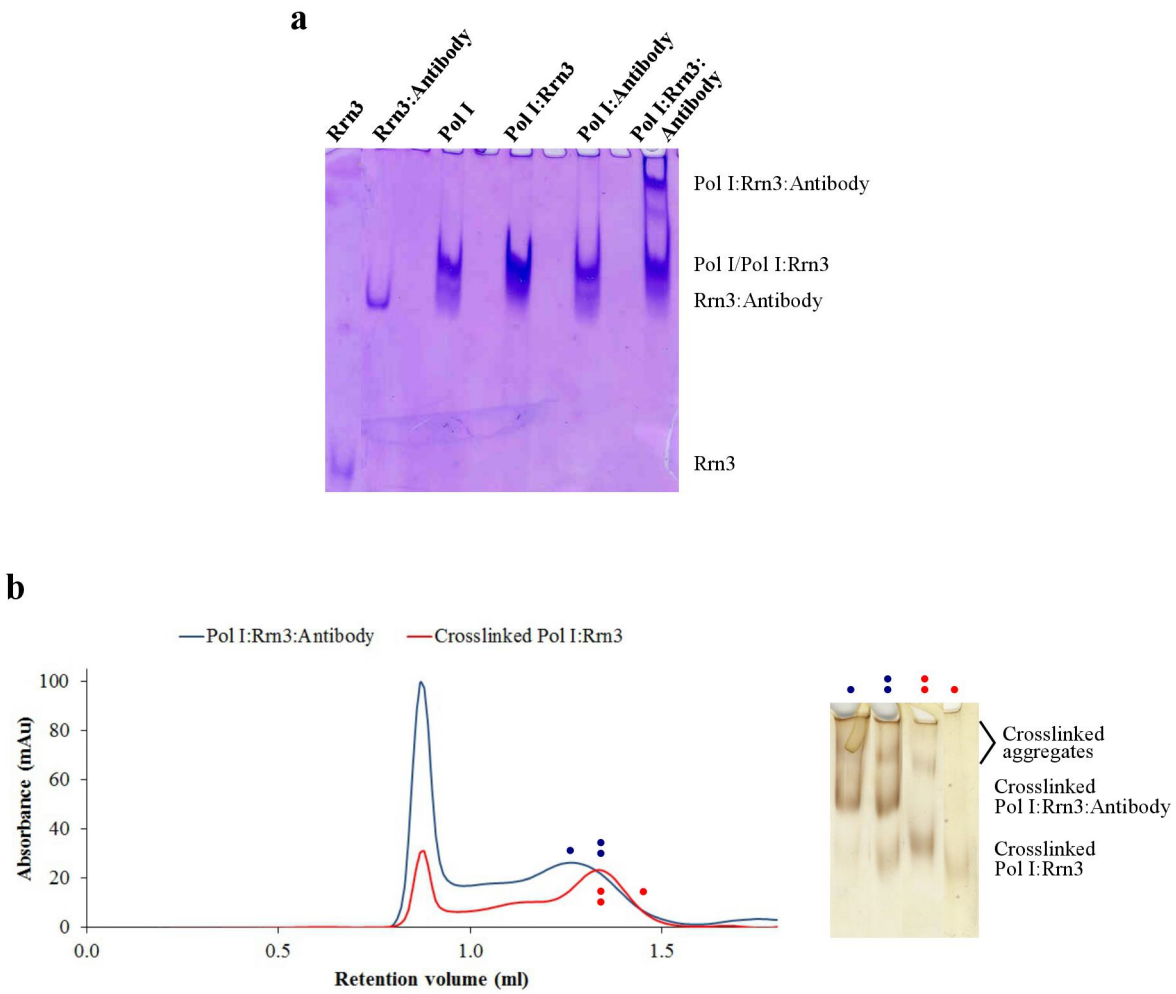

Figure 11. ***In vitro*** **assembly of Pol I:Rrn3:Antibody.**
(a) Pol I:Rrn3:Antibody interaction analysed in a coomassie-stained 4% native gel with controls. (b) Pol I:Rrn3 was crosslinked, incubated with the antibody and Pol I:Rrn3:Antibody run in a gel filtration. Crosslinked Pol I:Rrn3 was loaded as a control. Only the Abs_{280nm} is represented in the chromatograms. Blue and red spheres identify the fractions (left) shown in the silver-stained 4% native gel (right) and correspond to Pol I:Rrn3:Antibody and crosslinked Pol I:Rrn3 chromatograms, respectively.

7. Negative-Staining Electron Microscopy Structures of Pol I, Pol I:Rrn3 and Pol I:Rrn3:Antibody Complexes

In negative-staining electron microscopy, the sample is easily prepared and the high contrast conferred by the stain allows the use of much less particles than in cryo-EM. In spite of the limited resolution, negative-staining is useful to assess sample homogeneity and stoichiometry in a short period of time. In fact, preliminary processing of an unfixed Pol I dataset showed a significant population of Pol I ΔA49/A34.5. Similarly, a preliminary negative-staining Pol I:Rrn3 reconstruction lacked density for Rrn3. Such results helped in the decision of crosslinking Pol I and Pol I:Rrn3. The negative-staining 2D averages and 3D structures of monomeric Pol I, Pol I:Rrn3 and Pol I:Rrn3:Antibody were obtained from independent datasets. After glutaraldehyde fixation, Pol I ΔA49/A34.5 population was reduced and a 3D reconstruction of Pol I:Rrn3 obtained.

Pol I, Pol I:Rrn3 and Pol I:Rrn3:Antibody micrographs showed a wide range of particle orientations and a degree of homogeneity suitable for EM studies (Figure 12a). Pol I 2D averages and 3D structure revealed the general features of the enzyme (Figure 12b and c). The stalk, the cleft and the A49/A34.5 heterodimer were easily recognizable and the fitting of Pol I atomic structure (pdb 4C3H) validated the reconstruction. Pol I:Rrn3 2D averages exhibited a new extra density near the stalk (Figure 12b, salmon arrow). The subtraction of Pol I map from Pol I:Rrn3 reconstruction (Figure 12c) also revealed a new density. However, in virtue of the limited resolution achieved using negative-staining, the fitting of the atomic structure of Rrn3 monomer admitted opposite directionality, one with the N-ter towards the stalk, agreeing with the crosslinking-MS model (Blattner et al 2011), and other rotated approximately 180°, placing the C-ter in the vicinity of the stalk. To confirm the extra density was Rrn3 responsibility and to elucidate its orientation, Pol I:Rrn3 was reconstituted using Rrn3 fused to a C-ter tag with affinity for an antibody. Upon incorporation of the antibody, negative-staining 2D averages showed the antibody sticking out (Figure 12b, yellow arrow) and confirmed the extra density in Pol I:Rrn3 was in fact Rrn3. On the other hand, Pol I:Rrn3:Antibody 3D reconstruction (Figure 12d) unambiguously demonstrated the Rrn3 C-ter is placed around the dock domain of Pol I facing the back of the enzyme and, therefore, the N-ter is placed next to the stalk. Based on the antibody labelling, Rrn3 crystal structure (Blattner et al 2011) was fitted in Rrn3 density (Figure 12e). The negative-staining EM data were consistent with a previous crosslinking-MS model (Blattner et al 2011) and the biochemical evidence on Pol I:Rrn3 interaction through A43 stalk subunit (Peyroche et al 2000).

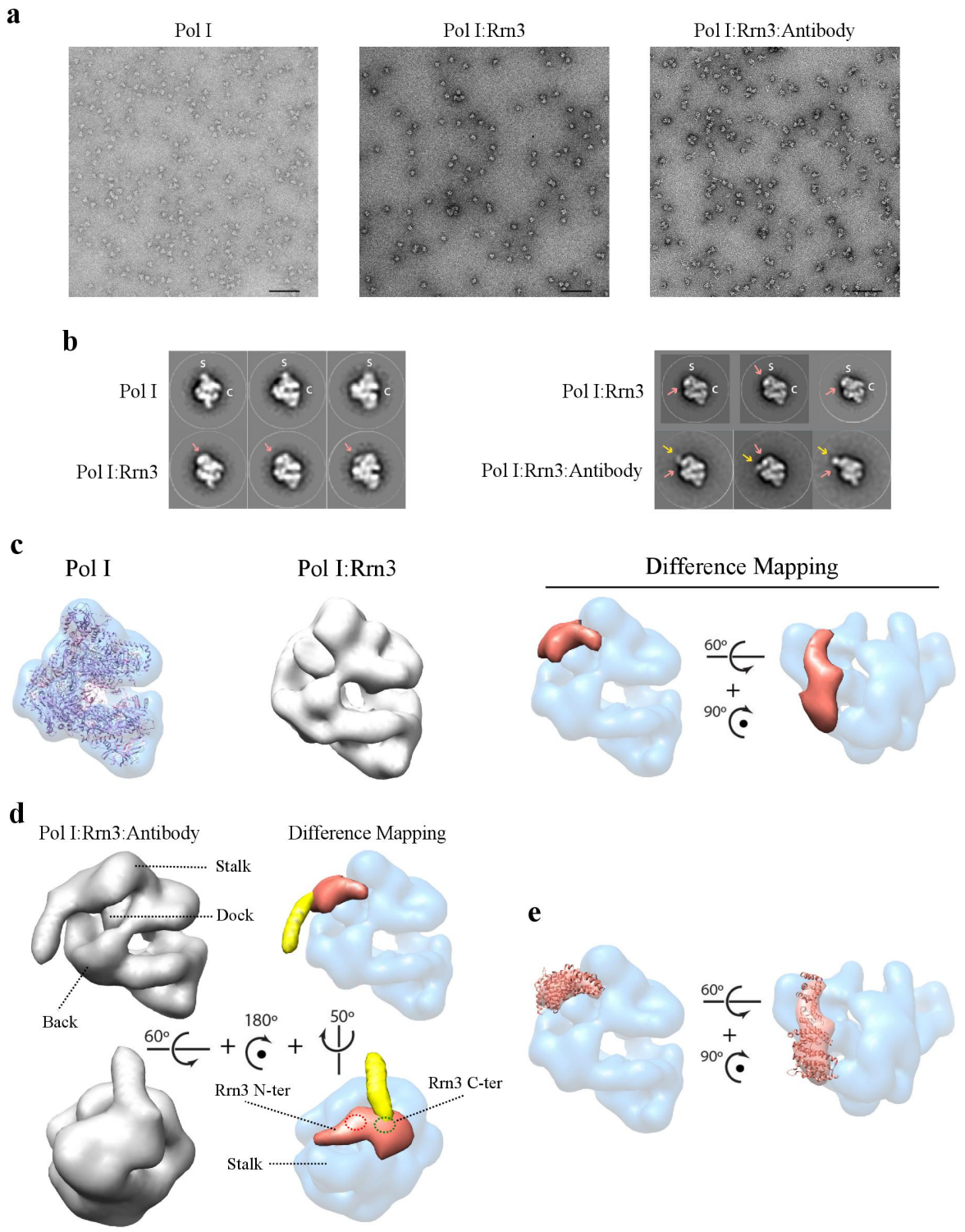

Figure 12. Negative-staining electron microscopy study of Pol I and Pol I:Rrn3.
(a) Typical field of negative-staining micrographs. Scale bars represent 100 nm. Samples crosslinked except the antibody in Pol I:Rrn3:Antibody. (b) Reference-free 2D averages of free monomeric Pol I, Pol I:Rrn3 and Pol I:Rrn3:Antibody. C, cleft; S, stalk. Salmon and yellow arrows indicate the position of Rrn3 and the Fab part of the antibody, respectively. (c) Negative-staining EM 3D reconstructions of free monomeric Pol I (Pol I pdb 4C3H fitted) and Pol I:Rrn3. Rrn3 density (salmon) obtained after difference mapping. (d) Negative-staining EM 3D reconstruction of Pol I:Rrn3:Antibody and Fab fragment density (yellow) obtained after difference mapping. Approximate locations of Rrn3 C-ter when fitting Rrn3 with its C-ter towards the stalk (dashed red ellipse) or the dock domain (dashed green ellipse) are indicated. The antibody density protrudes from the green region. (e) Fitting of the atomic structure of Rrn3 (3TJ1) based on the antibody labelling.

8. Cryo-EM Structures of Pol I and Pol I:Rrn3 Complexes

Pol I and Pol I:Rrn3 cryo-EM information was derived from the same dataset. The 2D micrographs revealed adequate protein concentration in the grid and a proper ice thickness (Figure 13a). Reference-free 2D averages showed a significant level of detail and in some cases the additional Rrn3 mass was visible (Figure 13b). 3D classification into 6 classes and using Pol I negative-staining reconstruction filtered at 60 Å as reference resolved monomeric Pol I (approximately 90000 particles) from Pol I:Rrn3 (approximately 32000 particles) (Figure 13c). In spite of glutaraldehyde fixation, 24200 particles, 13% of the particles in the 3D classification input, exhibited poor density for A49/A34.5 heterodimer. The percentage of particles without Rrn3 was much higher, around 80%. Separate 3D refinement of Pol I and Pol I:Rrn3 led to 5.6 and 7.7 Å structures, respectively (Figure 13c). Local resolutions in some Pol I:Rrn3 map regions reached 6.0 Å, allowing recognition of all α-helices. Since increasing the number of particles may improve resolution, Pol I and Pol I:Rrn3 particle sets were joined (new stack with 122500 particles) and subjected to iterative 3D refinement without Rrn3 masking and using Pol I filtered at 60 Å as reference. After per-particle motion correction (particle polishing in RELION) and post-processing map sharpening, the procedure yielded a map of monomeric Pol I identical to the 5.6 Å reconstruction but attaining a resolution of 4.9 Å (Figures 13c and 14a). Structural details of significant regions in monomeric Pol I corroborate the quality of the reconstruction and of the derived atomic model (Figure 14b). Fitting of Pol I atomic structure (4C3H) in the Pol I map at 4.9 Å (Figure 14c) evidenced relevant conformational changes in the transition from dimer to monomer, as the crystal structure was partially outside monomeric Pol I density, especially the stalk. It was also immediately perceptible that the cleft should close to fit monomeric Pol I map.

The cryo-EM map of Pol I:Rrn3 showed the location of the activating factor on the enzyme (Figure 15a). The thinner N-ter half of Rrn3 rests in a hollow formed by the stalk, A135 stalk-binding and A190 dock. The thicker C-ter half extends from the cavity where the N-ter fits towards the AC40/AC19 heterodimer, contacting the assembly platform. In spite of the average resolution being limited to 7.7 Å, Pol I:Rrn3 map quality (Figure 15b) allowed building of a pseudo-atomic model by rigid-body real-space refinement. Fitting of crystallographic Pol I structure (4C3H) in Pol I:Rrn3 map showed the stalk is fixed (Figure 15c). In comparison with the fitting of the same molecule in monomeric Pol I (Figure 14c), there was enough room in the 3D map to accommodate A43/A14 sub-complex. However, approximately the same regions of Pol I (4C3H) were outside the 3D Pol I and Pol I:Rrn3 envelopes, suggesting the major conformational changes occur in the transition from dimers to monomers rather than upon Rrn3 binding to Pol I (Figures 14c and 15c).

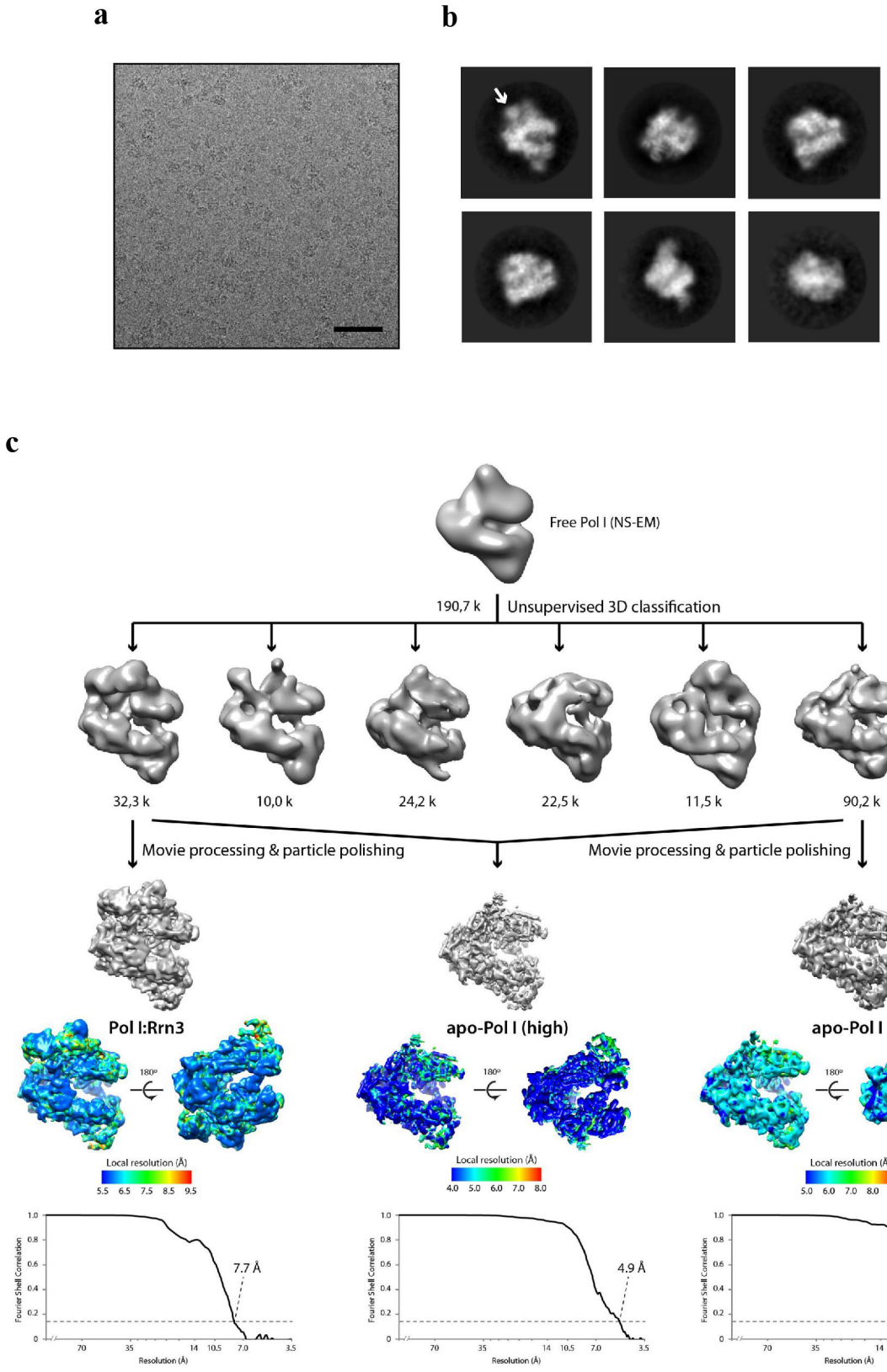

Figure 13. Cryo-EM processing of Pol I:Rrn3 dataset.
Extracted from Torreira et al 2017. (a) Typical field of crosslinked Pol I:Rrn3 cryo-EM grids. The scale bar represents 50 nm. (b) Initial reference-free 2D averages. An arrow indicates the position of Rrn3. (c) Data processing strategy showing the initial volume, the result of 3D classification and the three refined maps with their corresponding local resolution estimation and FSC curves.

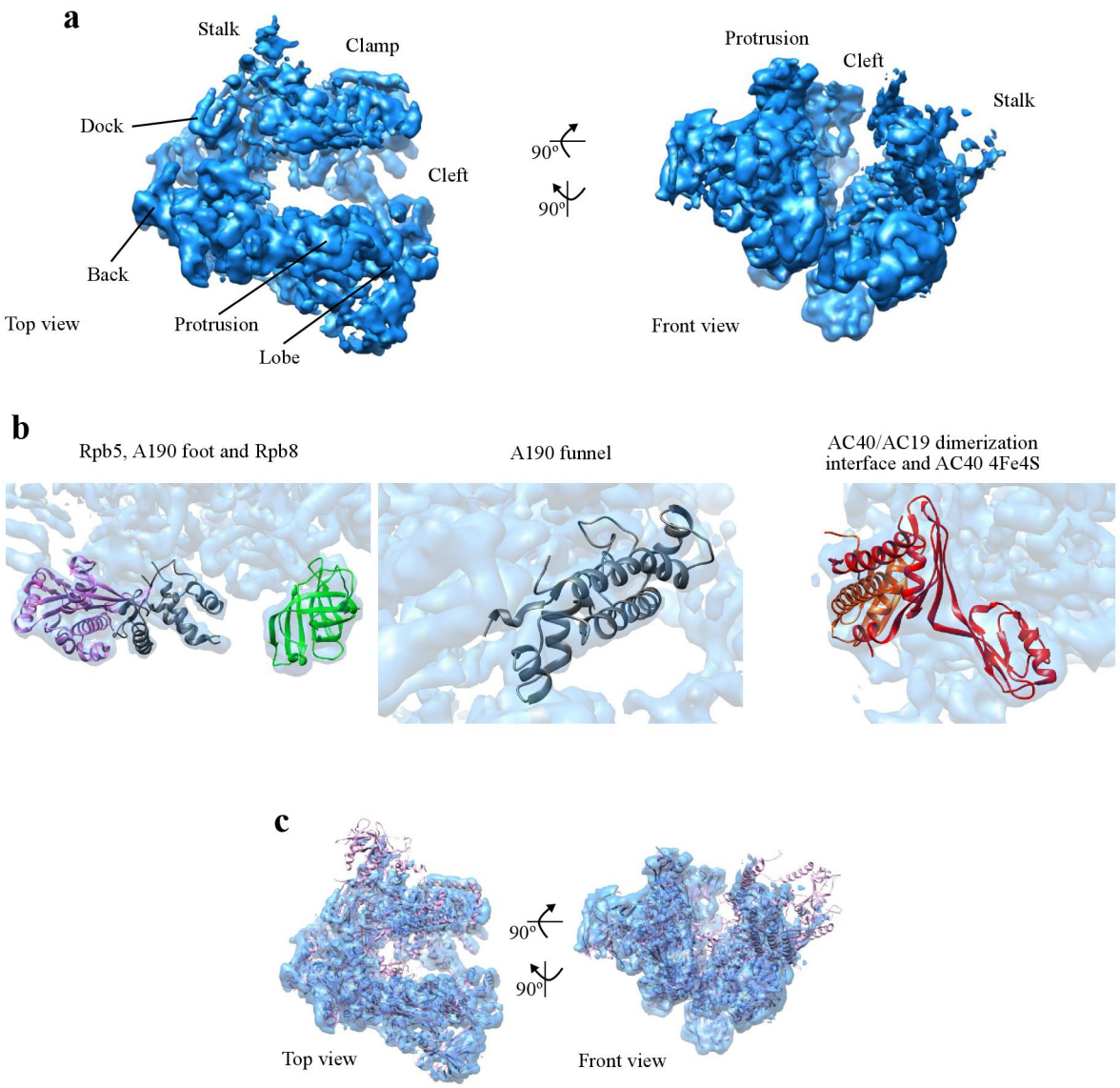

Figure 14. Cryo-EM structure of Pol I.
(a) Cryo-EM reconstruction of Pol I at 4.9 Å resolution. (b) Close-up views of regions in the 3D map of Pol I and selected domains of the pseudo-atomic model coloured as in Figure 4. (c) X-ray crystal structure of dimeric Pol I (4C3H) fitted in monomeric Pol I map.

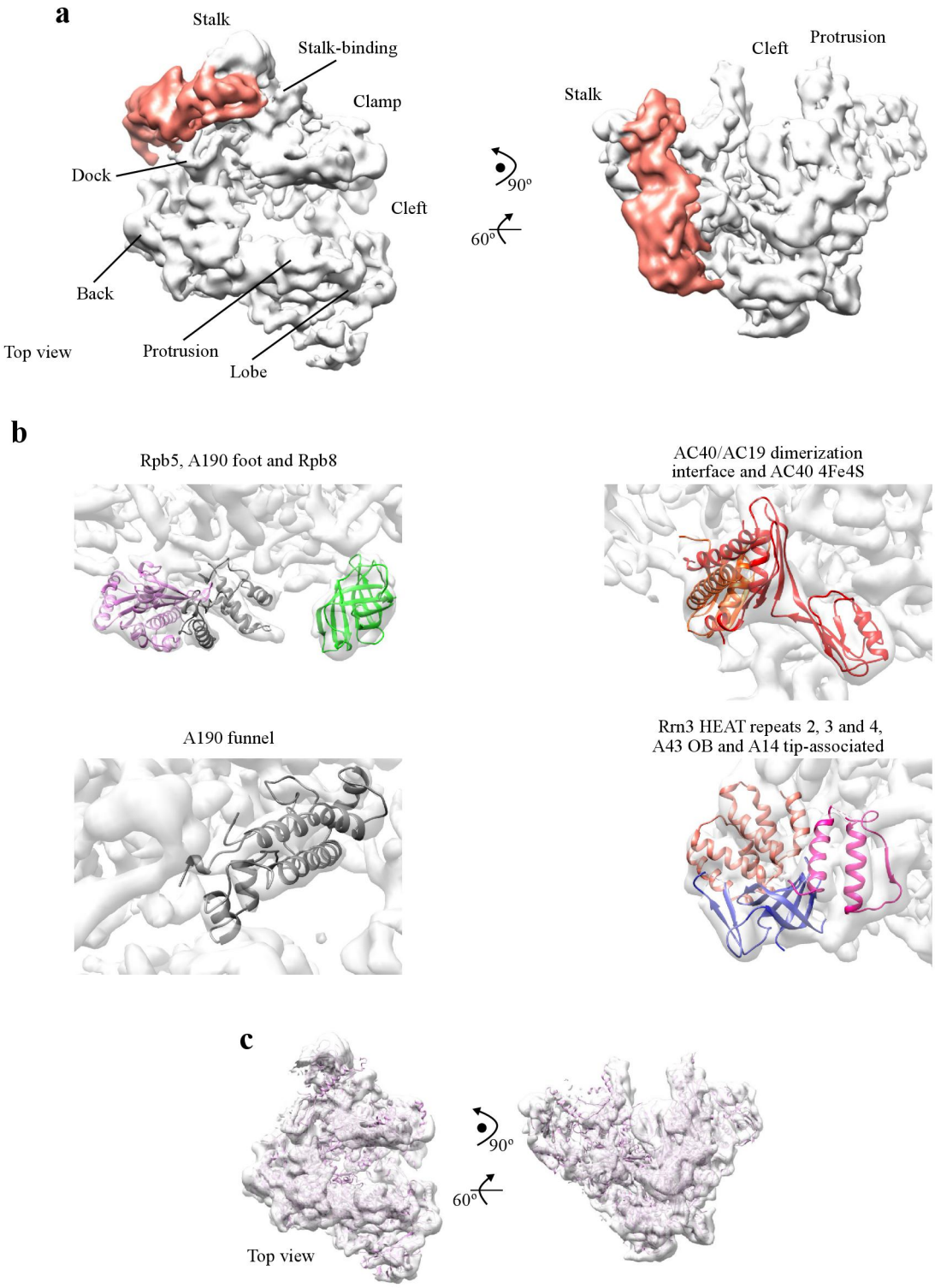

Figure 15. Cryo-EM structure of Pol I:Rrn3.
(a) Cryo-EM reconstruction of Pol I:Rrn3 at 7.7 Å resolution. Rrn3 in salmon. (b) Close-up views of regions in the 3D map of Pol I:Rrn3 and selected domains of the pseudo-atomic model coloured as in Figure 4. Rrn3 in salmon. (c) Pol I (4C3H) fitted in Pol I:Rrn3 map (Rrn3 density removed for clarity).

9. Pol I Conformational Changes in the Dimer to Monomer Transition

Fitting of the crystal structure of Pol I (4C3H) into the cryo-EM map of monomeric Pol I, deletion of the disordered regions and rigid-body real-space refinement provided an atomic model for free Pol I and an opportunity to follow dimer to monomer conformational changes (Figure 16a and Video 1. Dimer to monomer structural changes in Pol I). In the cryo-EM map, there is only weak density for the tip domain of A43 that contacts the enzyme core and for a short segment of the tip-associated (TA) domain in A14 (Figure 16a). Therefore, the stalk is highly flexible in free monomeric Pol I (Figures 16a, b and c). A43 C-ter displacement was predictable in the absence of a neighbour Pol I molecule stabilizing this region. On the contrary, the mobility of the whole OB domain in A43 was somehow surprising, as well as of a portion of the A135 stalk-binding domain (residues 1148 to 1156) (Figure 16b). Three additional smaller regions lack density in monomeric Pol I structure, being disordered: A12.2 C-ter Zn-ribbon, the central part of A190 bridge helix and the A190 DNA-mimicking loop (Figure 16b). The A49/A34.5 heterodimerization module is dissociable as 3D classification sorted out 14-subunit Pol I from Pol I ΔA49/A34.5. Both in Pol I cryo-EM monomer and crystallographic dimer, the same A49/A34.5 regions were modelled, missing A49 linker and A49tWH domain.

Apart from new flexible regions, Pol I dimer to monomer transition implies conformational changes (Figure 16c). In the dimer there is a 42 Å distance between the clamp and the protrusion whereas in the model of monomeric Pol I that distance is 38 Å (Figure 16c). Superimposition of the two models shows this partial closure of the cleft of approximately 4 Å. Protrusion in A135 and clamp coiled-coil in A190 approach each other, mainly through a movement of the last motif. The overall conformational changes also involve concomitant clamp head swing, movement of A190 jaw towards the clamp and opening of Rpb5 subunit and associated A190 region (Figure 16c).

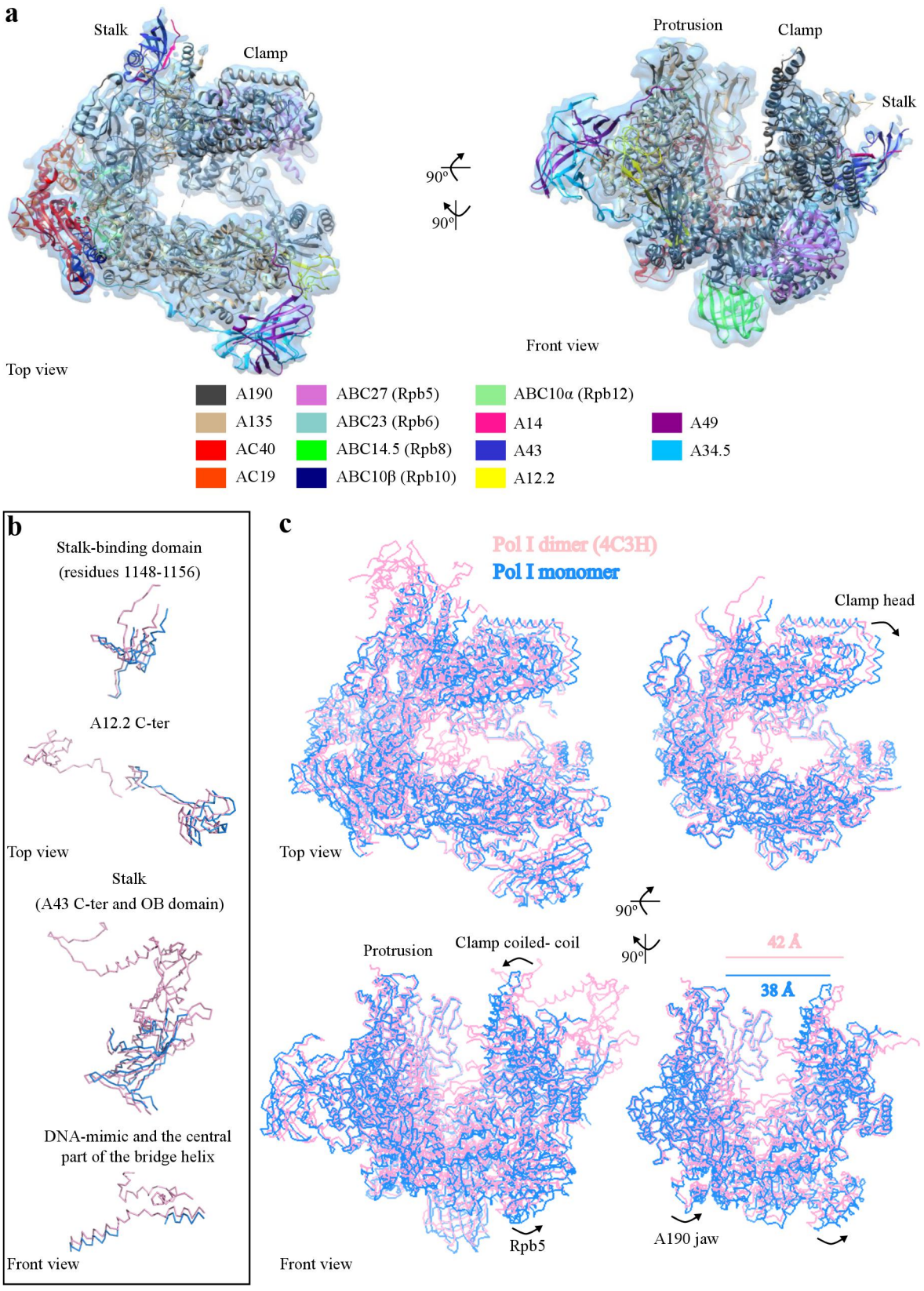

Figure 16. Dimer to monomer structural changes in Pol I.
(a) Cryo-EM reconstruction of Pol I at 4.9 Å resolution superposed with the derived pseudo-atomic model. (b) Close-up views of regions becoming flexible in the transition from dimeric (pink) to monomeric (blue) Pol I. (c) Comparison between the structural models of Pol I dimer (4C3H) and Pol I monomer after superimposition onto A135 (left). For clarity, only subunits A190 and A135 are shown (right). Arrows stress specific conformational changes.

10. Pol I:Rrn3 Interaction

Map interpretation was supported by fitting of the atomic structures of Pol I and Rrn3 followed by rigid-body refinement (Figure 17a). Rrn3 contacts at least 6 Pol I subunits (Figure 17b). The tightest interaction is established between Rrn3 N-ter and A43, as this subunit provides the broader binding surface in Pol I. Rrn3 N-ter further contacts a small region of A135 stalk binding domain. Pol I stalk interaction with Rrn3 also involves A14. The central part of the factor associates with A190 dock domain and, towards its C-ter, Rrn3 binds AC19/AC40 heterodimer.

OB domain of A43 strongly binds Rrn3 HEAT repeats 2, 3 and 4 (Figure 17b, upper left). In line with this, mutations in the central part of A43 abolished Pol I:Rrn3 complex formation (Peyroche et al 2000). HEAT repeats 2, 3 and 4 harbour the serine patch, in particular S145, in which a phosphomimetic mutation (S145D) impairs Pol I:Rrn3 interaction *in vitro* and causes a severe growth phenotype (Blattner et al 2011). Interestingly, S145 is in close proximity to a quadrilateral whose vertices are formed by residues S141, S143, S156 and S244 in the OB domain (Figure 17b, bottom left). The second stalk subunit, A14, contacts Rrn3 around HEAT repeat 5 using helix α2 of its TA domain (Figure 17b, upper left). Residues in this helix that are most proximal to Rrn3 include three consecutive serines (S83-85) and arginine 91 (R91). The TA-loop in A14, not visible in Pol I crystallographic structures, might be placed near Rrn3 and play a role in Rrn3 binding, despite also mobile in the cryo-EM Pol I:Rrn3 reconstruction. A135 stalk-binding domain, through residues in the region between 1118 and 1124, corresponding to a Pol I insertion not found in other multi-subunit RNAPs, binds the N-ter part of Rrn3 around HEAT repeat 3 (Figure17b, upper right). The main interaction between A190 and Rrn3 involves the α-helix formed between residues 550 and 560 in A190 dock domain and α-helix 10 of HEAT repeat 5 in Rrn3 (Figures 17b, upper right). The dock domain helix is also Pol I-specific. The disordered acidic loop between HEAT repeats 5 and 6 in Rrn3 occupies a position prone to bind Pol I, but new density in Pol I:Rrn3 map correspondent to the stabilization of this loop was not detected. Towards the back of Pol I, contact points are established between the C-ter third of Rrn3 with the N-ter of AC19 and the AC40 C-ter (Figure 17b, bottom right).

The reconstruction agrees with the described interaction between Rrn3 and A43 (Peyroche et al 2000) and with a crosslink between Rrn3 residue K558 to residue K329 in Pol I subunit AC40 (Blattner et al 2011). An additional crosslink between K558 in Rrn3 and K582 in A190 (Blattner et al 2011) is more difficult to rationalize but could still happen due to the mobility of the 21-amino acid loop after Rrn3 HEAT repeat 10 where the lysine stands.

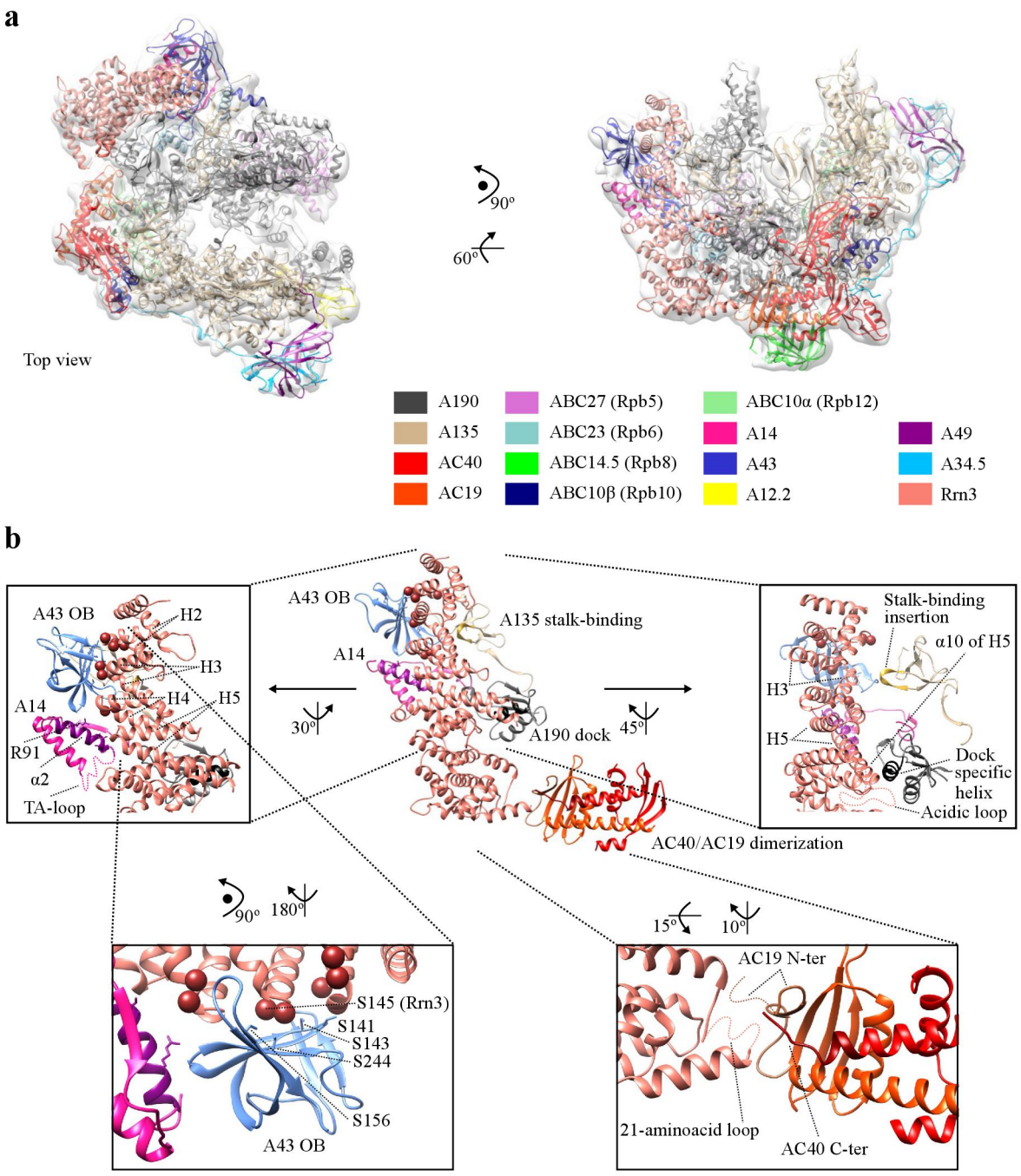

Figure 17. Pol I:Rrn3 interaction.

(a) Cryo-EM reconstruction of Pol I:Rrn3 at 7.7 Å resolution superposed with the derived pseudo-atomic model. (b) Details of Pol I:Rrn3 interaction. The view represented in the centre is the same as in (a) after rotation. Light blue, A43 OB; deep pink, A14; dark magenta, α2-helix of A14; salmon, Rrn3; brown spheres, serine patch; tan, A135 stalk-binding; goldenrod, Pol I-specific insertion between residues 1118 and 1124 in A135 stalk-binding; grey, A190 dock; black, Pol I-specific helix between residues 550 and 560 in A190 dock; orange, AC19; red, AC40 dimerization domain; light brown, residues 42 to 59 in AC19 N-ter; dark red, residues 326 to 335 in AC40 dimerization corresponding to the C-ter of AC40; dashed salmon, unresolved acidic loop and 21-amino acid loop (containing K558 crosslinked with A190 and AC40) in Rrn3 crystal structure (Blattner et al 2011) and Pol I:Rrn3 reconstruction; dashed deep pink, unresolved A14 TA-loop in Pol I crystal structures (Engel et al 2013, Fernández-Tornero et al 2013) and Pol I:Rrn3 reconstruction; dashed light brown, unresolved N-ter of AC19 in the same Pol I structures and Pol I:Rrn3 reconstruction. HEAT repeats 2, 3, 4 and 5 are indicated, as well as α10-helix in HEAT repeat 5 and specific serines in A43 OB or Rrn3.

11. Pol I and Rrn3 Structural Changes upon Pol I:Rrn3 Association

Pol I association with Rrn3 leads to the recovery of the stalk density found in the dimer, with exception of the A43 C-ter sequence (Figure 18). Binding of Rrn3 to Pol I fixes A14, A43 N-ter helix and A43 OB. Apart from stalk fixation, the structure of the enzyme in the Pol I:Rrn3 complex is identical to free monomeric Pol I. Thus, Pol I conformational changes in the transition from an inactive dimeric enzyme to an activated form in complex with Rrn3, able to be recruited to the rDNA promoter, are mainly due to monomerization rather than an effect of Rrn3 binding. In what respects to Rrn3, rigid body refinement indicated a minor bend in the C-ter (not shown in Figure 18) if compared to the X-ray crystal structure (Blattner et al 2011). Rrn3 C-ter approaches slightly the N-ter to contact AC40/AC19 heterodimer. None of the mobile regions in Rrn3 crystal structure (Blattner et al 2011) generated densities in the 3D cryo-EM reconstruction of Pol I:Rrn3 at 7.7 Å.

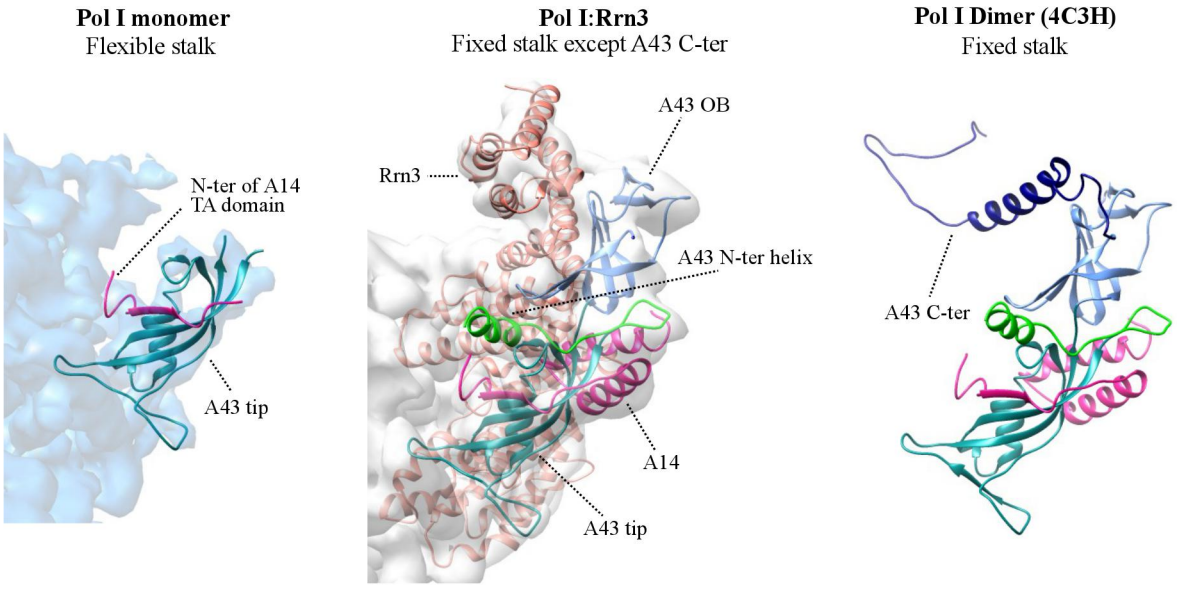

Figure 18. Stalk in Pol I monomer, Rrn3-bound Pol I and Pol I dimers.
Parts of Pol I and Pol I:Rrn3 maps are represented in blue and white transparencies, respectively. Pol I dimer stalk from Fernández-Tornero et al 2013.

DISCUSSION

In this work, Pol I and Rrn3 were separately expressed and purified. After incubation and glutaraldehyde fixation, cryo-EM Pol I:Rrn3 and Pol I structures were obtained from the same dataset. The resolution achieved allowed pseudo-atomic model building and comparison between inactive dimers, free monomeric and Rrn3-bound Pol I. Dimer to monomer transition implies flexibility in the stalk (A43 C-ter, A43 OB, A43 N-ter helix and most of A14), in a short region of A135 stalk-binding domain, A12.2 C-ter, DNA-mimicking loop and partially in the bridge helix. It also implies partial cleft closure and global interconnected conformational changes. Free monomeric Pol I is not transcriptionally competent as the bridge helix is mobile at the central region and cleft contraction is insufficient to trap downstream template DNA. Despite incomplete cleft closure, Rrn3-bound Pol I corresponds to the activated state of the enzyme in the sense that it can be recruited to the growing PIC and subsequently initiate transcription. Rrn3 stabilizes monomeric Pol I through stalk fixation (except A43 C-ter) and avoids Pol I homodimerization (steric hindrance to the binding of a neighbour enzyme). Rrn3 and the stabilized stalk may form a surface for interactions with promoter-bound TFs. Main structural changes occur in the dimer to monomer transition and are conserved upon Rrn3 binding. Also in this work, sedimentation velocity assays suggest homodimeric Rrn3 might establish a dimer-monomer equilibrium dependent on its concentration. At last, EMSAs indicate yeast Rrn3 might not be a DNA-binding protein, at least independently of TFs or in the absence of post-translational modifications.

1. The Road to Pol I:Rrn3 Assembly and the Importance of Chemical Fixation

Yeast Pol I was endogenously expressed in large fermenters and initially purified taking advantage of a TAP-tag C-terminally fused to AC40 (Moreno-Morcillo et al 2014) or A190. After anion-exchange chromatography, complete Pol I was reproducibly isolated, although with a minor presence of Pol I ΔA49/A34.5, and resolved from Pol I A43ΔCt, Pol I A43ΔCt ΔA49/A34.5 and most of Pol I ΔA49/A34.5. Yeast Rrn3 was overexpressed in *E.coli* and purified by nickel affinity, anion-exchange and gel filtration chromatographies. *In vitro* incubation of Pol I and Rrn3 with subsequent fixation generated Pol I:Rrn3 suitable for negative-staining and cryo-EM studies.

The fact that yeast Rrn3 was expressed in bacteria did not impede Pol I:Rrn3 complex formation, indicating that post-translational modifications are not required for the association. This is in agreement with a previous Pol I:Rrn3 assembly *in vitro* (Blattner et al 2011) and bacterial co-expression and interaction of A43:Rrn3 (Peyroche et al 2000) but differs from mouse Rrn3, that if produced in bacteria does not interact with A43 *in vitro* (Cavanaugh et al 2002). Possibly, post-translational modifications in Rrn3 are required for Pol I:Rrn3 assembly in higher eukaryotes but not in yeast. Regarding Pol I, its phosphorylation is required to form a stable and transcriptionally

active Pol I:Rrn3 complex (Fath et al 2001). In this work, Pol I purified either in the presence or absence of phosphatase inhibitors interacted with Rrn3, possibly meaning that the isolation protocol without phosphatase inhibitors still yields the enzyme with the required phosphorylation(s) for Rrn3 association.

Pol I A49/A34.5 contains a lobe-binding dimerization module conserved in Pol II initiation factor TFIIF and Pol III C53/C37 (Carter and Drouin 2010, Geiger et al 2010, Vannini and Cramer 2012). A49/A34.5 also presents the A49tWH, a DNA-binding domain with structural and functional homology to Pol II initiation factor TFIIE (Geiger et al 2010, Pilsl et al 2016 a, Tafur et al 2016). Pol I purification customarily yielded a secondary Pol I ΔA49/A34.5 population and the protocol was not able to completely resolve 14-subunit from 12-subunit Pol I. Furthermore, it is not clear if pure 14-subunit Pol I can establish equilibrium with 12-subunit enzyme lacking A49/A34.5 heterodimer or if the dissociation is simply result of complex instability *in vitro*. It would be interesting to further study this subject *in vitro* and perhaps conceive strategies to evaluate the presence of the free heterodimer and Pol I ΔA49/A34.5 *in vivo*.

The binding of Pol I and Rrn3 is not a complete reaction. It was shown that free Rrn3 is present in the 1:1 molar incubation and most particles in Pol I:Rrn3 EM datasets did not contain Rrn3, even after chemical crosslinking. Furthermore, Pol I itself was demonstrated to be a source of heterogeneity, as even the samples most enriched in 14-subunit Pol I contained Pol I ΔA49/A34.5 enzyme. Most likely, the early choice for cryo-EM as a technique to obtain detailed structural information on the activated state of Pol I was right. Sample heterogeneity would have probably been an obstacle for Pol I:Rrn3 crystallization whereas it could be resolved by the 2D and 3D classification processing protocols. Moreover, even the attempts to crystallize Pol I ΔA49/A34.5:Rrn3 or Pol I A43ΔCt ΔA49/A34.5:Rrn3 were not successful.

Glutaraldehyde crosslinking was useful to fix Pol I:Rrn3, as a preliminary reconstruction without fixation lacked density for the activating factor. The risk of protein distortion can be almost completely ruled out because Pol I and Pol I:Rrn3 *artificial* deformities were not visible in the 3D structures and, most importantly, Pol I:Rrn3 reconstruction was essentially identical to two published structures (Engel et al 2016, Pilsl et al 2016 a), only with minor differences. All three reconstructions of activated Pol I (Engel et al 2016, Pilsl et al 2016 a, Torreira et al 2017) used chemically fixation in sample preparation, further confirming the incompleteness of the binding reaction between Pol I and Rrn3 and Pol I as heterogeneity source.

StrepMAB-Immo (IBA), presented as an antibody for capturing of Strep-tag® II fusion proteins on solid phases (IBA web page), was successfully used to label Rrn3Strep2 in the crosslinked Pol I:Rrn3 complex. The assembly of Pol I:Rrn3:Antibody complex constitutes the unique example reported to date of the utilisation of this concrete antibody to label a protein for EM studies.

Equally noteworthy, the Fab fragment of the antibody was detected in negative-staining reference-free 2D averages and unambiguously in the Pol I:Rrn3:Antibody 3D reconstruction when adding the antibody to the beforehand fixed Pol I:Rrn3. Conceivably, this strategy avoided the "collapse" of the antibody against Pol I:Rrn3, occurring in the initial approach of assembling the ternary complex and only then chemically crosslink the sample.

2. Cryo-EM Structure of Pol I:Rrn3 as an Opportunity to Investigate the Role of Pol I Domains

The fact that Pol I and Pol I:Rrn3 structures were derived from a single dataset attests the potential of cryo-EM to analyse distinct protein populations inside the same sample. Sample heterogeneity is now becoming an opportunity rather than an obstacle (Nogales and Scheres 2015), much because 3D classification strategies emerged (Scheres et al 2007, Sigworth et al 2010, Scheres 2016) capable of sorting out related but not identical protein complexes.

In the Pol I:Rrn3 cryo-EM structure presented here, HEAT repeats 2, 3 and 4 form a surface that strongly interacts with A43 subunit. Rrn3 serine patch lies in HEAT repeats 2, 3 and 4 and contains S145 and S185, shown to be important for Pol I:Rrn3 interaction (Blattner et al 2011). In mammals, phosphorylation of the serine patch represses Pol I transcription (Mayer et al 2004, Mayer et al 2005), presumably preventing Pol I interaction with Rrn3 counterpart TIF-IA (Engel et al 2016). Notably, Rrn3 S145 sits at the centre of a quadrilateral whose vertices are formed by A43 residues S141, S143, S156 and S244 in OB domain. A possible effect of mutations in these A43 residues, separately or together, has not been investigated. In *S. cerevisiae*, a global phosphorylation analysis after DNA damage identified S244 of A43 as being phosphorylated (Albuquerque et al 2008). Thus, it could be interesting to investigate the growth phenotype generated by mutations in S244 and in the other serines forming the quadrilateral, as well as their effect in Pol I:Rrn3 interaction.

Since Rrn3 around HEAT repeat 5 contacts helix α2 of A14 TA domain, A14 mutants were generated and their *in vivo* effects evaluated (Torreira et al 2017). Pol I:Rrn3 interaction was studied using PICT (Protein interactions from Imaging of Complexes after Translocation) (Gallego et al 2013, Torreira et al 2017), a fluorescence microscopy technique in living cells, whereas Pol I and Rrn3 promoter occupation was assessed by ChIP. A14 deletion caused a 75% decrease in Pol I:Rrn3 levels inside the cell and similar reduction in promoter association for Pol I and Rrn3. A strain lacking A14 C-ter tail (A14ΔCt, Δ101-137), flexible residues in the cryo-EM structures presented here and Pol I crystal structures but that could lie near Rrn3, presented similar levels to wild-type cells in Pol I:Rrn3 complex and promoter association, indicating the tail is not involved in Rrn3 binding. Nevertheless, a strain also lacking helix α2 (A14Δα2Ct, Δ80-137) presented about

one third reduction in detected Pol I:Rrn3 and promoter occupation by the enzyme and initiation factor. Moreover, a point mutation in R91 (R91E), residue in α2 and in close proximity to Rrn3, induced coincident outcomes. Thus, in addition to A43 role in Rrn3 binding (Peyroche, Milkereit et al. 2000), the cryo-EM Pol I structure and the mutational analysis demonstrated A14 is crucial for Pol I:Rrn3 interaction, in particular the α2 of A14 TA domain (Torreira et al 2017). Like R91, a stretch of three serines (S83-S85) in α2 of A14 is in close proximity to Rrn3. It would be interesting to perform a mutational analysis also in these residues.

An internal loop in A14 TA domain (TA-loop, residues 51 to 79), situated near Rrn3 and like A14 C-ter tail disordered in crystal and cryo-EM structures, could influence binding of the transcription factor to Pol I. TA-loop was truncated (A14ΔTAloop) and the mentioned techniques used to analyse Pol I:Rrn3 interactions in living cells and promoter occupation (Torreira et al 2017). Strikingly, A14ΔTAloop strain presented a 2-fold increase in detected Pol I:Rrn3 complexes in relation to wild type yeast, as well as increased Pol I and Rrn3 promoter association. This result suggests A14 TA-loop counteracts Rrn3 binding to Pol I and, therefore, enzyme recruitment to the rDNA promoter. It remains unclear how and in which circumstances A14 TA-loop opposes Rrn3 association to Pol I or, in a scenario mimicked by A14ΔTAloop, if the loop can be "deactivated" to favour the interaction.

3. Pol I Structural Changes Occur Mainly upon Monomerization

The most significant Pol I structural changes occur in the dimer to monomer transition rather than upon Rrn3 binding. In monomeric Pol I, the stalk becomes highly mobile and such flexibility possibly matters for Rrn3 recognition and binding in the course of Pol I activation. A12.2 C-ter Zn-ribbon domain, containing the catalytic loop responsible for RNA cleavage in backtracked Pol I transcribing complex, is also disordered. The cleavage activity is not required until nucleotide incorporation begins and from that point of view the stabilization of A12.2 C-ter is not necessary in free monomeric Pol I. The bridge helix, close to the three aspartic acid residues and divalent metal ion in the catalytic centre, conserved among multi-subunit RNAPs (Weinzierl 2011) and key element promoting RNAP activity (Kaplan and Kornberg 2008), is unwound in the central region in dimeric inactive Pol I crystal structures (Engel et al 2013, Fernández-Tornero et al 2013). Its central part is disordered in free monomeric Pol I, as the corresponding electron density lacks. Therefore, A190 bridge helix must fold completely in a transcriptionally active state. Additionally, in monomeric Pol I, DNA-mimicking loop flexibility in the cleft, in contrast to most inactive dimeric X-ray crystal structures, constitutes a step towards an enzyme able to transcribe because the template DNA can now access the active site.

There are global and interconnected conformational changes upon Pol I monomerization distinct from the aforementioned new flexible regions provoking density loss in the 3D map. The roughly 4 Å cleft closure in relation to the open cleft in Pol I inactive dimers does not make Pol I ready to transcribe but represents a movement towards the transcribing state. In fact, the structure of a transcribing Pol I complex (Tafur et al 2016) detected a 31 Å cleft opening and, therefore, a 11 Å closure compared to dimeric Pol I, which presented a 42 Å distance between the clamp and the protrusion (Fernández-Tornero et al 2013). Furthermore, comparisons based on an independent elongation complex estimated the cleft closure from inactive dimer to transcribing Pol I being up to 13 Å (Neyer et al 2016). Finally, the conformational changes in Pol I upon monomerization are not result of tension loosening from packed molecules in the crystal because a cryo-EM 3D structure of Pol I dimers in solution (Pilsl et al 2016 a) resembled the crystallographic structures (Engel et al 2013, Fernández-Tornero et al 2013).

4. Similar Overall Conclusions Drawn from Independent Cryo-EM Structures of Monomeric and Rrn3-Bound Pol I

The proteins in cryo-EM are preserved in a quasi-native state much like in solution. X-ray and cryo-EM models from inactive dimeric Pol I strongly resemble (Engel et al 2013, Fernández-Tornero et al 2013, Pilsl et al 2016 a). They differ only slightly in the position of the DNA-mimicking loop (also referred to as expander) inside the cleft, although a particular Pol I crystallographic form of inactive dimers (pdb 4C3I, Fernández-Tornero et al 2013) presents the DNA-mimicking loop almost completely disordered and displays poor density for A12.2 C-ter Zn-ribbon (Moreno-Morcillo et al 2014), indicating flexibility of these motifs. Still, crystal structures can be used to draw conclusions on the structural changes happening in solution transitioning from Pol I dimers to monomers. Including the one presented here, three independent cryo-EM structures of free monomeric Pol I at 7.5 Å (Pilsl et al 2016 a), 4.0 Å (Neyer et al 2016) and 4.9 Å (Torreira et al 2017) have been recently reported, allowing comparison with Pol I dimers and between free monomers. Three independent cryo-EM structures of Pol I:Rrn3 complexes at 7.5 Å (Pilsl et al 2016 a), 4.8 Å (Engel et al 2016) and 7.7 Å (Torreira et al 2017) were published as well. Hence, 3D changes upon Rrn3 association to Pol I and the details of Pol I:Rrn3 interaction could be evaluated in distinct laboratories.

A12.2 C-ter Zn-ribbon, responsible for intrinsic Pol I RNA cleavage activity, was shown to be flexible in two monomeric Pol I reconstructions and two Pol I:Rrn3 structures (Pilsl et al 2016 a, Torreira et al 2017). The correspondent cryo-EM models from one particular laboratory (Neyer et al 2016) displayed most of A12.2 C-ter Zn-ribbon stabilized, although A12.2 C-ter catalytic hairpin was also mobile (Engel et al 2016). In transcriptionally inactive dimers (Engel et al 2013,

Fernández-Tornero et al 2013, Pilsl et al 2016 a), A12.2 C-ter Zn-ribbon emerged most of the times inside the NTP entry pore, placing the catalytic hairpin next to the active site, but in some cases also disordered. It seems that A12.2, although a "built-in" transcription factor, can be removed from Pol I active centre, or at least partially disordered, to switch off RNA cleavage ability if pyrophosphorolysis is not required. A similar phenomenon was described for C11 subunit, A12.2 counterpart in Pol III (Hoffmann et al 2015). Pol II, in turn, relies on an independent and dissociable elongation and cleavage factor, TFIIS, to remodel the catalytic site and confer RNA cleavage activity to the enzyme (Jeon et al 1994, Gu and Reines 1995, Kettenberger et al 2003). Therefore, TFIIS is homologous to the catalytic C-ter Zn ribbon of A12.2 in Pol I and C11 in Pol III (Vannini and Cramer 2012).

A43 C-ter (also referred to as connector), imperative for Pol I dimerization in crystals and solution (Engel et al 2013, Fernández-Tornero et al 2013, Pilsl et al 2016 a), becomes mobile in all monomeric and Rrn3-bound Pol I reconstructions. With respect to A43 OB, it is found highly disordered in free monomeric Pol I. In spite of this, the laboratory that released monomeric Pol I cryo-EM map at 4.0 Å (Neyer et al 2016) applied a mask around A43/A14 and allowed only local searches during refinement (focused refinement in the stalk), this way modelling A43 OB even with poor density. Pol I stalk flexibility is higher in free monomers, with strong density only for A43 tip and a segment of A14 TA domain. A43 OB becomes visible upon Rrn3 binding. Thus, Rrn3 induces stalk fixation stabilizing Pol I monomer and at the same time stereochemically impedes Pol I dimers formation. However, Rrn3 does not induce A43 C-ter fixation as it happens in Pol I homodimerization, where the stalk is almost entirely fixed and A43 C-ter plays its crucial dimerization role.

The DNA-mimicking loop appeared mobile in the three Pol I:Rrn3 structures (Engel et al 2016, Pilsl et al 2016 a, Torreira et al 2017) and in the higher resolution monomeric Pol I reconstructions (Neyer et al 2016, Torreira et al 2017). Displacement of this A190 loop, obstructing the active site in most Pol I dimers, is a requisite to DNA entry in the cleft and transformation of inactive Pol I in transcription-competent enzyme. Such structural change most likely occurs upon Pol I monomerization and is not related to Rrn3 activation. Regarding A190 bridge helix, critically influencing catalysis and DNA translocation, it is partially unwound in inactive dimers and mobile in its central region in Pol I monomers or Rrn3-bound Pol I. However, in recently reported transcribing complexes (Neyer et al 2016, Tafur et al 2016), the totality of Pol I bridge helix folds into a conformation similar to other multi-subunit RNAP ECs. Therefore, while DNA-mimicking loop removal from the cleft might take place upon monomerization, bridge helix folding probably occurs concomitantly to promoter direct physical association. Monomeric Pol I and activated Pol I are non-transcribing states of the enzyme.

In all monomeric and Rrn3-bound Pol I reconstructions the cleft was shown to contract in relation to dimeric Pol I. Moreover, binding of the activating factor to Pol I monomers did not contribute to further cleft closing. Nonetheless, in the higher resolution structures (Neyer et al 2016, Torreira et al 2017) a partial closure was detected while in the 7.5 Å reconstructions (Pilsl et al 2016 a) the movement was broader, nearly the same experienced from Pol I dimers to Pol I transcribing complexes (Neyer et al 2016, Tafur et al 2016). Explanations for these discrepancies remain elusive but may reside in distinct experimental conditions that trapped different intermediary Pol I states. In summary, monomeric or activated Pol I, in the presence of template DNA, would have to suffer structural changes in order to transcribe. On one hand, transcription is favoured in these forms as Pol I is not dimeric, the DNA mimicking-loop is not obstructing the active site and A12.2 C-ter catalytic hairpin is not in place to change enzyme`s ribonucleotide polymerization activity to RNA cleavage. On the other hand, A190 bridge helix is not completely folded and the higher resolution structures displayed only partial cleft contraction, features incompatible with transcription.

In what respects to Rrn3 in Pol I:Rrn3 reconstructions, the initiation factor is positioned almost identically in all cases and the majority of the described contacts between Pol I and the activating factor coincide. Rrn3 “serine patch” is oriented towards Pol I and residues in this patch and HEAT repeats 2 to 4 participate in the strong interaction with the OB domain in A43. An interaction around S83-85 in A14 and E224 in Rrn3 is detected (Pilsl et al 2016 a, Torreira et al 2017) and most likely happens in the higher resolution 3D reconstruction (Engel et al 2016). A contact between a Pol I-specific insertion in A135 stalk-binding domain and the N-ter part of Rrn3 is also identified (Engel et al 2016, Torreira et al 2017), probably happening in the 7.5 A structure (Pilsl et al 2016 a). All three Pol I:Rrn3 complexes present a contact surface between Rrn3 helix-forming residues starting around E243, before the disordered acidic loop, and the dock domain of A190, mainly through a Pol I-specific insertion. The three works refer identical contacts between the α-helix in Rrn3 immediately before the 21-amino acid flexible loop and the AC40/AC19 sub-complex. An additional contact with Rpb6 subunit is mentioned in the highest resolution Pol I:Rrn3 structure (Engel et al 2016). All main Pol I conformational changes observed in the dimer to monomer transition are retained in the respective Pol I:Rrn3 maps. The major Pol I structural change upon Rrn3 association is stalk stabilization (A43 OB fixation).

5. Pol I:Rrn3 Association is Specific but Parallels with Other Transcription Systems Can Be Established

No homologous factor to eukaryotic Rrn3 has been found in archaea or bacteria. In addition, Rrn3 requirement for Pol I recruitment to the rDNA promoter in a potential productive state constitutes a special case among eukaryotic RNAPs and aRNAP, which can bind promoter-bound GTFs and do

not rely on a TF for recruitment. bRNAP also behaves differently because, although demanding σ like Pol I needs Rrn3 for recruitment, the holoenzyme (bRNAP:σ) triggers RNA synthesis without involvement of extra promoter-bound factors. By the above, the interaction between Pol I and Rrn3 was expected to be specific and the Pol I:Rrn3 structures recently published (Engel et al 2016, Pilsl et al 2016 a, Torreira et al 2017) confirmed that uniqueness. Rrn3 binds a Pol I surface that no other TF in nuclear, bacterial or archaeal RNAPs entirely contacts in their correspondent enzyme. In particular, Pol I-specific insertions in the A135 stalk-binding domain and the A190 dock domain may exist only for contact Rrn3 and might constitute features that evolved to meet the specific function of Rrn3 association.

However, topological similarities between Rrn3 position in Pol I:Rrn3 and the Mediator* in a Pol II:Mediator complex using a 15-subunit core Mediator comprising all its essential subunits (Plaschka et al 2015) can be identified. Like Rrn3 in Pol I, the coactivator* contacts Pol II at the stalk, the dock domain in Rpb1 (homologous to A190) and the Rpb3/Rpb11 heterodimer (homologous to AC40/AC19). The similarity in the contact is particularly evident looking at Med18/Med20 (especially Med18) interaction with Rpb3/Rpb11 sub-complex and Pol II dock domain (Torreira et al 2017).

In comparison with the bacterial system, a topological resemblance can also be identified between σ domain 4 and Rrn3 (Torreira et al 2017). Fitting of the bacterial holoenzyme structure (pdb 1IW7) into the Pol I:Rrn3 cryo-EM map presented in this work shows σ domain 4 and Rrn3 bind equivalent regions of the RNAP, the β-flap and the dock domain, respectively.

6. Pol I Homodimers and Pol I:Rrn3 Participate in the Regulation of rDNA Transcription

In parallel with the structural biology of yeast Pol I activation, Carlos Fernández-Tornero's laboratory searched for collaborations to investigate the regulation of rDNA transcription through Pol I:Rrn3 and Pol I:Pol I complexes, in particular after growth arrest or growth restoration. In Oriol Gallego's laboratory, *in vivo* levels of Rrn3-bound Pol I and enzyme homodimers were evaluated using PICT. In Olga Calvo's laboratory, Pol I and Rrn3 recruitment to the promoter were measured by ChIP.

It was shown that growth arrest either by nutrient deprivation, inhibition of rRNA maturation or inhibition of protein synthesis induces the formation of Pol I dimers, not significantly detected in growing cells (Torreira et al 2017). Presumably, the dimers observed *in vivo* are the same kind previously reported *in vitro* (Milkereit et al 1997, Engel et al 2013, Fernández-Tornero et al 2013, Pilsl et al 2016 a). Such dimers inhibit PIC formation occluding the binding sites for Rrn3 and CF.

It was further demonstrated that nutrient deprivation is followed by Pol I:Rrn3 clearance and, later, the assembly of a pool of Pol I dimers. Replenishment of nutrient supply first induces Pol I:Rrn3 association, then Pol I:Pol I disruption. Worthy of note, Pol I:Rrn3 levels are affected in the first place in both situations while enzyme dimerization inside the cell is initiated or disrupted more slowly. Rrn3 functions to recruit Pol I to rDNA promoter in an activated state (Pol I:Rrn3) potentially productive, as Pol I recruited alone is not able to initiate transcription. Pol I homodimers, in turn, could protect a pool of the highly abundant enzyme from degradation in growth arrest conditions, saving energy to the cell. Also, once favourable growth circumstances are regained, Pol I:Pol I disruption could help in the full recovery of 35S pre-rRNA synthesis, avoiding *de novo* Pol I production, especially if few ribosomes are available.

The essential role of A43 C-ter in Pol I homodimerization, suggested by the crystal structures (Engel et al 2013, Fernández-Tornero et al 2013), was confirmed *in vivo* (Torreira et al 2017). A mutation in A43 C-ter, partially truncating the β-hairpin and deleting the acidic C-ter tail (A43Δ307-326) but conserving the α-helix, was enough to block Pol I dimers formation under starvation in yeast. Besides, it was shown Pol I:Rrn3 interaction is not compromised in this Pol I mutant.

A role beyond storage of Pol I enzymes and energy saving was proposed for Pol I homodimerization in the same work (Torreira et al 2017). Combining PICT and ChIP analyses in wild type and Pol I mutant yeast strains, a wealth of data suggested that Pol I homodimerization likewise participates in Pol I transcription inactivation resulting from nutrient deprivation. A model in which Pol I:Rrn3 and Pol I:Pol I complexes modulate rDNA transcription was proposed. When nutrients are depleted, Pol I:Rrn3 levels and promoter association drop exponentially. Pol I homodimerizes subsequently and, therefore, initial transcription inactivation is mainly driven by Pol I:Rrn3 disassembly whereas Pol I homodimerization limits rDNA transcription at a later stage. Upon refeeding from starvation, the available Pol I:Rrn3 complexes are rapidly recruited for transcription and new Pol I:Rrn3 assemblies may form. Disruption of Pol I homodimers possibly constitutes a later response mechanism that provides fresh monomeric Pol I able to interact with Rrn3 and further activate rDNA transcription.

Finally, PICT analysis suggested yeast Pol II and Pol III do not dimerize in growing or nutrient depletion conditions, although experimental constraints in the case of Pol II should lead to careful data interpretation (Torreira et al 2017). Homodimerization likely is a special feature of Pol I and not a general mechanism to regulate eukaryotic transcription. However, it cannot be excluded that Pol II and Pol III homodimerize in situations other than nutrient starvation.

7. New Structures of the Core Factor Are in Line with this Work

While writing this thesis, the X-ray crystal structure of the CF at 3.2 Å and the cryo-EM structures of Pol I:Rrn3:CF at 7.7 Å and of a Pol I initially transcribing complex (Pol I ITC, Pol I:Rrn3:CF:DNA/RNA scaffold) at 3.4 Å were reported (Engel et al 2017). X-ray crystallographic structure of the CF reveals a bimodular heterotrimer. Module I is formed by the N-ter and the β-propeller of Rrn6 and subunit Rrn11, with two helical domains around Rrn6. Module II is formed by most of Rrn7 and the helical C-ter of Rrn6 wrapped around the two cyclin domains of Rrn7. Two regions protrude from Rrn7, a helix from an insertion in the C-ter cyclin domain that reaches module I and 93 unsolved N-ter residues containing a Zn-ribbon domain. The CF assumed different positions in Pol I:Rrn3:CF and Pol I ITC, although employing the same three regions to bind Pol I:Rrn3. In both cases, CF bound activated Pol I through subunit Rrn11 in module I, the protruding cyclin insertion helix in Rrn7 and the Rrn7 N-ter ribbon, mobile in the crystal structure. The association between Rrn7 N-ter ribbon and both A190 dock and a Rrn3 loop between HEAT repeats 4 and 5 was the only interaction kept in Pol I:Rrn3:CF and Pol I ITC. The other two CF regions involved in Pol I:Rrn3 recruitment contacted Pol I but not Rrn3 using distinct Pol I surfaces in Pol I:Rrn3:CF and Pol I ITC. Importantly, a previously described interaction between Rrn3 and Rrn6 subunit of the CF (Peyroche et al 2000) was not detected in Pol I:Rrn3:CF or in the Pol I initially transcribing complex (Engel et al 2017).

The *in vivo* relevance of a Pol I:Rrn3:CF complex is not confirmed, as the CF is thought to recognize the rDNA promoter and only then bind activated Pol I. Nevertheless, in Pol I:Rrn3:CF reconstruction, Pol I and Rrn3 are found as in the reported Pol I:Rrn3 structures, meaning that the activating factor is identically positioned and the enzyme displays partial cleft closure. Thus, as it happens in Pol I:Rrn3, Pol I:Rrn3:CF stabilizes the enzyme in an intermediate state not transcriptionally active. Also, one module of the bimodular CF structure contacts the upstream end of the cleft, blocking promoter DNA loading into the active site. Actually, although using the same regions to bind Pol I:Rrn3, the CF drastically moved in Pol I ITC to the outer surface of Pol I, allowing DNA loading and cleft contraction.

Stalk fixation is the major conformational change in Pol I upon Rrn3 binding. The same way stalk flexibility might be relevant for specific Rrn3 association, it may also be thought that stalk stabilization generates a new surface together with Rrn3 for the interaction with promoter-bound initiation factors. In fact, cryo-EM structures of Pol I:Rrn3:CF and Pol I ITC showed a contact between the N-ter ribbon domain of Rrn7 subunit of CF and Rrn3 loop between HEAT repeats 4 and 5 (Engel et al 2017). The N-ter ribbon responsible for Rrn3 contact also binds A190 dock and mutations in this Rrn7 region had a drastic effect in Pol I initiation (Engel et al 2017), suggesting

that these particular CF-Pol I and CF-Rrn3 interactions may be critical in rDNA transcription. Also, Pol I stalk fixation by Rrn3 is maintained through these CF-bound states without additional flexibility or new A43/A14 density.

In the recent Pol I ITC cryo-EM structure, no interaction between promoter DNA (around position -35 and +20 relative to TSS) and Rrn3 was found, as the initiation factor is positioned far from the nucleic acid scaffold (Engel et al 2017). This is in agreement with the EMSA presented here, in which Rrn3 did not bind a 100 bps rDNA fragment spanning the TSS. Furthermore, contrary to rat Rrn3 (Stepanchick et al 2013), a DNA-binding domain in yeast Rrn3 was not described to date. However, yeast Rrn3 has been expressed in *E.coli* (Blattner et al 2011, Engel et al 2016, Pilsl et al 2016 a, Engel et al 2017, Torreira et al 2017) and post-translational modifications could confer DNA-binding capacity to the initiation factor, although the distances involved make such a scenario unlikely. There is also the possibility that Rrn3 association with rDNA depends on proteins other than CF and Pol I subunits, although a direct physical interaction was observed between mammalian Rrn3 and a rDNA fragment (Stepanchick et al 2013). Therefore, it seems yeast Rrn3 does not bind DNA in Pol I PIC, in spite of high Rrn3 conservation in eukaryotes. In addition, human Rrn3 was found to be a GTP-binding protein and that property important in regulation of rRNA synthesis in T cells (Nguyen et al 2015). Therefore, mechanistic differences in PIC assembly between yeast and mammalian may exist, as well as distinct mechanisms of transcription initiation regulation.

Pol I:Rrn3, Pol I:Rrn3:CF (Engel et al 2017) and a Pol I OC (Tafur et al 2016) exhibit a partial cleft closure, intermediate between inactive dimers and transcribing complexes. Furthermore, in spite of DNA-mimicking loop displacement and A12.2 C-ter catalytic hairpin mobility, these structures display a bridge helix incompletely folded. Thus, in Pol I complexes lacking a DNA/RNA scaffold, the enzyme does not assume a state compatible with ribonucleotide polymerisation. In turn, Pol I ITC or ECs (Neyer et al 2016, Tafur et al 2016, Engel et al 2017) display full cleft contraction appropriate for transcription and the A190 bridge helix completely folded. These Pol I complexes also present the active site not remodelled by A12.2 and a clear cleft, as A12.2 C-ter and the DNA-mimicking loop are mobile. Hence, Pol I only seems to achieve a productive conformation upon Rrn3 binding, promoter recruitment and immediately before incorporation of the first ribonucleotide, as all the required Pol I transformations are not finished in the OC.

8. Future Perspectives on Pol I Initiation

Rrn3 forms a dimer in solution but binds Pol I as a monomer (Blattner et al 2011). While at concentrations of 2 mg/mL or higher Rrn3 seems to be mainly dimer, ultracentrifugation data suggest monomer-dimer equilibrium *in vitro* is favoured upon further dilution and at 0.2 mg/mL it is established. It is possible that *in vivo* the nuclear and nucleolar Rrn3 concentrations are far below this value and in such conditions a monomeric Rrn3 population is free to bind and activate Pol I. This hypothesis is further supported by estimations of Rrn3 copy number in logarithmic growth yeast cells ranging from 23 (Kulak et al 2014) to 138 ± 29.5 molecules per cell (Ghaemmaghami et al 2003). However, it seems also plausible that in the nucleolus or other locations inside the cell Rrn3 concentrations keep elevated, forcing the Pol I activating factor to dimerize. In this case, a so far unknown signal would trigger Rrn3 dimers disruption so the monomer could associate with Pol I. Therefore, investigating Rrn3 oligomerisation states *in vivo* constitutes a challenge. Although the current PICT assay (Gallego et al 2013, Torreira et al 2017) requires a more abundant protein, like any of Pol I subunits, in order to be effective, other methods must be experimented.

As with Rrn3, the precise cellular mechanisms triggering Pol I homodimerization or dimers disruption remain elusive. Dimer break is not triggered by Rrn3, as incubating *in vitro* up to a 10-fold excess of Rrn3 with dimeric Pol I caused only a subtle shift towards monomers compared to controls without Rrn3 (Pilsl et al 2016 a). DNA-binding is also unlikely because the cleft and the active site are inaccessible to the promoter. Other TF than Rrn3 or post-translational modifications, especially involving kinases or phosphatases, or even the action of small molecules, could play a role in Pol I dimers assembly or disruption.

In this work, the systematic Pol I A43ΔCt co-purification during Pol I isolation, as well as the result from the western blot detection of A43 over the growth curve, suggest a Pol I pool missing A43 C-ter might exist in living yeast cells. This should be further investigated, as well as the hypothetical Pol I A43ΔCt increase in stationary phase. Interestingly, Pol I homodimers would protect A43 C-ter from proteolytic cleavage. Moreover, Pol I A43ΔCt, lacking last 49 C-ter residues, was shown to bind Rrn3 *in vitro* similarly to Pol I and a mutation in A43 C-ter partially truncating the β-hairpin and deleting the acidic C-ter tail (A43Δ307-326, deleting 20 residues) did not affect Pol I:Rrn3 levels intracellularly (Torreira et al 2017). Thus, it remains to be determined if the truncated enzyme plays a role in the regulation of rDNA transcription.

Cryo-EM structures of Pol I:Rrn3 show details of the interaction and structural changes in the enzyme. However, the reconstructions do not explain how Pol I:Rrn3 complex is disrupted in starvation or the signals behind its assembly upon growth restoration. In yeast, Pol I phosphorylation seems required to form a stable Pol I:Rrn3 complex (Fath et al 2001) and 5 subunits, including A43, present phosphosites (Gerber et al 2008). However, the phosphorylated

residue or combination of residues important for Rrn3 binding was not identified. Also in yeast, there is no record of a specific Rrn3 phosphorylation favouring or repressing Pol I:Rrn3 interaction *in vivo*. *In vitro*, two phosphomimetic mutations in Rrn3 serine patch (S145D and S185D) strongly reduced the binding of the activating factor to Pol I (Blattner et al 2011). The effect on Pol I:Rrn3 assembly of specific Rrn3 phosphorylations is better known in humans. Phosphorylation of S199 or T200, correspondent to S185/S186 of the serine patch in yeast Rrn3, impairs Pol I:Rrn3 formation with a negative effect in rDNA transcription (Mayer et al 2004, Mayer et al 2005). S170/S172 phosphorylation and dephosphorylation, in turn, are coordinated to, respectively, trigger the release of Rrn3 from Pol I after initiation or facilitate their association, sustaining multiple rounds of transcription (Bierhoff et al 2008). Both modifications have a positive effect in rRNA synthesis. Other described phosphorylations and dephosphorylations, either activating or repressing human Rrn3, were not implicated in the assembly or disruption of Pol I:Rrn3. In summary, further research to understand how Pol I:Rrn3 levels are modified to regulate Pol I promoter association and rRNA synthesis is required.

The new Pol I:Rrn3:CF and Pol I ITC structures suggest the CF recruits Pol I:Rrn3 to the rDNA promoter mainly through Pol I contacts. The importance of the single CF-Rrn3 contact must be assessed. *In vitro* transcription on immobilized templates indicated Pol I can be recruited to the promoter without Rrn3, yielding a PIC transcriptionally inactive (Aprikian et al 2001). Similarly, in mammals, Pol I can be recruited to a template without TIF-IA but is unable to catalyse the formation of the first phosphodiester bond (Schnapp et al 1993, Cavanaugh et al 2008). In the light of current knowledge, free monomeric Pol I recruitment can be explained but the reasons for transcription failure remain unknown. Is Pol I stalk fixation by Rrn3 essential for initiation? Is Rrn3 critical for Pol I recruitment to the promoter or simply stabilizes monomeric Pol I impeding its homodimerization? Is there a post-recruitment role for Rrn3?

In addition, Rrn3 is not part of the elongating Pol I complex (Milkereit and Tschochner 1998, Bier et al 2004). Strains lacking Pol I A49 subunit or its last 28 amino acids showed strongly reduced release of Rrn3 during elongation in ChIP assays (Beckouet et al 2008), suggesting a role of the conserved C-ter WH domain of A49 (A49tWH) in Rrn3 dissociation from the transcribing complex. Therefore, further research is essential to know how and when the activating factor is displaced from the enzyme in transcription initiation or early elongation.

In conclusion, the specific requirement for Rrn3 in Pol I transcription initiation is not yet fully understood, especially knowing the other eukaryotic and archaeal multi-subunit RNAPs do not rely on such an initiation factor and bind the promoter-bound TFs by themselves. And also being aware that bRNAP uses σ for enzyme recruitment but σ is a multifunctional TF sufficient to trigger bRNAP initiation.

CONCLUSIONS / CONCLUSIONES

CONCLUSIONS

1. Cryo-EM structures of monomeric Pol I and Rrn3-bound Pol I showed structural changes in the transition of the enzyme from inactive dimers to monomers and from these to the activated state in complex with Rrn3.

2. Most significant structural changes occur in the dimer to monomer transition. The majority of Pol I stalk becomes flexible, as well as A12.2 C-ter Zn-ribbon, the central part of the bridge helix and the DNA-mimicking loop. Also, there are global conformational changes implying partial cleft closure towards an intermediate state still incompatible with ribonucleotide polymerisation.

3. Rrn3 binding to Pol I fixes the stalk (except the C-ter of A43, crucial for homodimerization) without other significant structural changes.

4. *In vitro* results suggest Rrn3 could establish a dimer-monomer equilibrium *in vivo* and that yeast Rrn3 is not a DNA-binding protein as the mammalian homolog.

CONCLUSIONES

1. Las estructuras de crio-microscopía electrónica de Pol I monomérica y Pol I unida a Rrn3 mostraron cambios estructurales en la transición de la enzima de dímeros inactivos a monómeros y de éstos al estado activado en complejo con Rrn3.

2. Los cambios estructurales más significativos ocurren en la transición de dímero a monómero. La mayoría del *stalk* de Pol I se vuelve flexible, así como el *A12.2 C-ter Zn ribbon*, la parte central de la *bridge helix* y el *DNA-mimicking loop*. Además, existen cambios conformacionales globales que implican un cierre parcial del *cleft* hacia un estado intermedio todavía incompatible con la polimerización de ribonucleótidos.

3. La unión de Rrn3 a Pol I fija el *stalk* (excepto el C-ter de A43, crucial para la homodimerización) sin otros cambios estructurales significativos.

4. Los resultados *in vitro* sugieren que Rrn3 podría establecer un equilibrio dímero-monómero *in vivo* y que Rrn3 de levadura no es una proteína de unión al DNA como la proteína homóloga en mamíferos.

GLOSSARY

Concept	Definition
Acrocentric Chromosomes	Those with the centromere very close to one end, creating two unequal arms. In humans, chromosomes 13, 14, 15, 21 and 22 are acrocentric chromosomes. The short arm contains repeating sequences such as the NORs.
Capping	First step in pre-mRNA processing in which a methylated guanosine is added to the 5`end of the nascent transcript. Capping occurs in early elongation.
Closed Complex (CC)	PIC before DNA opening.
Coactivator	TF that participates in transcription activation but does not bind the DNA itself, rather binding to one or more activators and relaying the information to GTFs.
General transcription factor (GTF)	TF that, *in vivo*, is used always or most often by a given RNAP. Each RNAP has its own GTFs that bind the core promoter, although, in eukaryotes, some subunits may be common to the transcription machinery of two or even three RNAPs. GTFs are many times basal transcription factors (BTFs), TFs strictly required by a given RNAP to drive *in vitro* transcription. All BTFs are GTFs but not all GTFs are BTFs.
HEAT repeat	Protein structural motif that consists of two antiparallel α-helices separated by a short loop. HEAT (Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A) and the yeast kinase TOR1).
Holoenzyme	Catalytically active enzyme consisting of an apoenzyme combined with its cofactor.
Initially transcribing complex (ITC)	PIC after DNA opening and with few ribonucleotides incorporated in the nascent RNA.
Mediator	Present in all eukaryotes, it is a large and conformationally flexible protein complex with a variable subunit composition. Participates in organization of chromatin architecture and regulation of Pol II pre-initiation, initiation, re-initiation, pausing and elongation (reviewed, among others, in Allen and Taatjes 2015, Hantsche and Cramer 2017).
Multi-subunit RNA polymerase	Any RNAP composed of more than one subunit found in Bacteria, Archaea and Eukarya. There are single-subunit RNAPs composed of a single polypeptide chain found in mitochondria, chloroplasts and many viruses like bacteriophage T7.
Open complex (OC)	PIC after DNA opening and before incorporating the first ribonucleotide.
Promoter	The set of DNA elements required for pre-initiation complex assembly and transcription initiation.
Splicing	Editing of the nascent precursor mRNA (pre-mRNA) in which introns are removed and exons covalently linked together.
TATA box	Conserved DNA sequence (most commonly TATAAA) located 25-35 bps before the TSS of many eukaryotic and archaeal genes that binds TBP or other TFs.
TOR	Target of rapamycin (TOR) is a highly conserved kinase in eukaryotes and a growth regulator that senses and integrates stimuli from changes in the levels of growth factors or amino acids, as well as in the energy of the cell or cellular stresses.

Concept	Definition
Transcription bubble	Structure formed in transcription in which the double-stranded DNA is unwound, fundamental to initiation, elongation and termination.
Transcriptional proofreading	Removal of a misincorporated ribonucleotide during transcription that involves RNAP backtracking and excision of a dinucleotide containing the misincorporation. RNA cleavage occurs at the same active site used for polymerization, contrary to DNA polymerases proofreading.

BIBLIOGRAPHY

Abrishami, V., J. Vargas, X. Li, Y. Cheng, R. Marabini, C. Sorzano and J. M. Carazo (2015). "Alignment of direct detection device micrographs using a robust Optical Flow approach." J Struct Biol **189**(3): 163-176.

Adams, P. D., P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-221.

Adrian, M., J. Dubochet, J. Lepault and A. W. McDowall (1984). "Cryo-electron microscopy of viruses." Nature **308**(5954): 32-36.

Albuquerque, C. P., M. B. Smolka, S. H. Payne, V. Bafna, J. Eng and H. Zhou (2008). "A multidimensional chromatography technology for in-depth phosphoproteome analysis." Mol Cell Proteomics **7**(7): 1389-1396.

Allegretti, M., D. J. Mills, G. McMullan, W. Kühlbrandt and J. Vonck (2014). "Atomic model of the F420-reducing [NiFe] hydrogenase by electron cryo-microscopy using a direct electron detector." Elife **3**: e01963.

Allen, B. L. and D. J. Taatjes (2015). "The Mediator complex: a central integrator of transcription." Nat Rev Mol Cell Biol **16**(3): 155-166.

Aprikian, P., B. Moorefield and R. H. Reeder (2001). "New model for the yeast RNA polymerase I transcription cycle." Mol Cell Biol **21**(15): 4847-4855.

Bai, X. C., I. S. Fernandez, G. McMullan and S. H. Scheres (2013). "Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles." Elife **2**: e00461.

Bazett-Jones, D. P., B. Leblanc, M. Herfort and T. Moss (1994). "Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF." Science **264**(5162): 1134-1137.

Beckmann, H., J. L. Chen, T. O'Brien and R. Tjian (1995). "Coactivator and promoter-selective properties of RNA polymerase I TAFs." Science **270**(5241): 1506-1509.

Beckouet, F., S. Labarre-Mariotte, B. Albert, Y. Imazawa, M. Werner, O. Gadal, Y. Nogi and P. Thuriaux (2008). "Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription." Mol Cell Biol **28**(5): 1596-1605.

Bedwell, G. J., F. D. Appling, S. J. Anderson and D. A. Schneider (2012). "Efficient transcription by RNA polymerase I using recombinant core factor." Gene **492**(1): 94-99.

Bell, S. P., R. M. Learned, H. M. Jantzen and R. Tjian (1988). "Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis." Science **241**(4870): 1192-1197.

Bier, M., S. Fath and H. Tschochner (2004). "The composition of the RNA polymerase I transcription machinery switches from initiation to elongation mode." FEBS Lett **564**(1-2): 41-46.

Bierhoff, H., M. Dundr, A. A. Michels and I. Grummt (2008). "Phosphorylation by casein kinase 2 facilitates rRNA gene transcription by promoting dissociation of TIF-IA from elongating RNA polymerase I." Mol Cell Biol **28**(16): 4988-4998.

Birch, J. L. and J. C. Zomerdijk (2008). "Structure and function of ribosomal RNA gene chromatin." Biochem Soc Trans **36**(Pt 4): 619-624.

Bischler, N., L. Brino, C. Carles, M. Riva, H. Tschochner, V. Mallouh and P. Schultz (2002). "Localization of the yeast RNA polymerase I-specific subunits." EMBO J **21**(15): 4136-4144.

Blattner, C., S. Jennebach, F. Herzog, A. Mayer, A. C. Cheung, G. Witte, K. Lorenzen, K. P. Hopfner, A. J. Heck, R. Aebersold and P. Cramer (2011). "Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth." Genes Dev **25**(19): 2093-2105.

Blombäck, B., M. Blombäck, P. Edman and B. Hessel (1966). "Human fibrinopeptides. Isolation, characterization and structure." Biochim Biophys Acta **115**(2): 371-396.

Bodem, J., G. Dobreva, U. Hoffmann-Rohrer, S. Iben, H. Zentgraf, H. Delius, M. Vingron and I. Grummt (2000). "TIF-IA, the factor mediating growth-dependent control of ribosomal RNA synthesis, is the mammalian homolog of yeast Rrn3p." EMBO Rep **1**(2): 171-175.

Bon, M., S. J. McGowan and P. R. Cook (2006). "Many expressed genes in bacteria and yeast are transcribed only once per cell cycle." FASEB J **20**(10): 1721-1723.

Bonissone, S., N. Gupta, M. Romine, R. A. Bradshaw and P. A. Pevzner (2013). "N-terminal protein processing: a comparative proteogenomic analysis." Mol Cell Proteomics **12**(1): 14-28.

Bordi, L., F. Cioci and G. Camilloni (2001). "In vivo binding and hierarchy of assembly of the yeast RNA polymerase I transcription factors." Mol Biol Cell **12**(3): 753-760.

Boukhgalter, B., M. Liu, A. Guo, M. Tripp, K. Tran, C. Huynh and L. Pape (2002). "Characterization of a fission yeast subunit of an RNA polymerase I essential transcription initiation factor, SpRrn7h/TAF(I)68, that bridges yeast and mammals: association with SpRrn11h and the core ribosomal RNA gene promoter." Gene **291**(1-2): 187-201.

Brilot, A. F., J. Z. Chen, A. Cheng, J. Pan, S. C. Harrison, C. S. Potter, B. Carragher, R. Henderson and N. Grigorieff (2012). "Beam-induced motion of vitrified specimen on holey carbon film." J Struct Biol **177**(3): 630-637.

Campbell, E. A., O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz and S. A. Darst (2002). "Structure of the bacterial RNA polymerase promoter specificity sigma subunit." Mol Cell **9**(3): 527-539.

Campbell, M. G., A. Cheng, A. F. Brilot, A. Moeller, D. Lyumkis, D. Veesler, J. Pan, S. C. Harrison, C. S. Potter, B. Carragher and N. Grigorieff (2012). "Movies of ice-embedded particles enhance resolution in electron cryo-microscopy." Structure **20**(11): 1823-1828.

Carrillo, E., G. Ben-Ari, J. Wildenhain, M. Tyers, D. Grammentz and T. A. Lee (2012). "Characterizing the roles of Met31 and Met32 in coordinating Met4-activated transcription in the absence of Met30." Mol Biol Cell **23**(10): 1928-1942.

Carter, R. and G. Drouin (2010). "The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors." Mol Biol Evol **27**(5): 1035-1043.

Cavanaugh, A. H., A. Evans and L. I. Rothblum (2008). "Mammalian Rrn3 is required for the formation of a transcription competent preinitiation complex containing RNA polymerase I." Gene Expr **14**(3): 131-147.

Cavanaugh, A. H., I. Hirschler-Laszkiewicz, Q. Hu, M. Dundr, T. Smink, T. Misteli and L. I. Rothblum (2002). "Rrn3 phosphorylation is a regulatory checkpoint for ribosome biogenesis." J Biol Chem **277**(30): 27423-27432.

Chen, D., A. S. Belmont and S. Huang (2004). "Upstream binding factor association induces large-scale chromatin decondensation." Proc Natl Acad Sci U S A **101**(42): 15106-15111.

Cipollina, C., J. van den Brink, P. Daran-Lapujade, J. T. Pronk, D. Porro and J. H. de Winde (2008). "Saccharomyces cerevisiae SFP1: at the crossroads of central metabolism and ribosome biogenesis." Microbiology **154**(Pt 6): 1686-1699.

Claypool, J. A., S. L. French, K. Johzuka, K. Eliason, L. Vu, J. A. Dodd, A. L. Beyer and M. Nomura (2004). "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." Mol Biol Cell **15**(2): 946-956.

Colland, F., N. Fujita, A. Ishihama and A. Kolb (2002). "The interaction between sigmaS, the stationary phase sigma factor, and the core enzyme of Escherichia coli RNA polymerase." Genes Cells **7**(3): 233-247.

Darst, S. A. (2001). "Bacterial RNA polymerase." Curr Opin Struct Biol **11**(2): 155-162.

de la Rosa-Trevín, J. M., J. Otón, R. Marabini, A. Zaldívar, J. Vargas, J. M. Carazo and C. O. Sorzano (2013). "Xmipp 3.0: an improved software suite for image processing in electron microscopy." J Struct Biol **184**(2): 321-328.

de la Rosa-Trevín, J. M., A. Quintana, L. Del Cano, A. Zaldívar, I. Foche, J. Gutiérrez, J. Gómez-Blanco, J. Burguet-Castell, J. Cuenca-Alba, V. Abrishami, J. Vargas, J. Otón, G. Sharov, J. L. Vilas, J. Navas, P. Conesa, M. Kazemi, R. Marabini, C. O. Sorzano and J. M. Carazo (2016). "Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy." J Struct Biol **195**(1): 93-99.

Denissov, S., M. van Driel, R. Voit, M. Hekkelman, T. Hulsen, N. Hernandez, I. Grummt, R. Wehrens and H. Stunnenberg (2007). "Identification of novel functional TBP-binding sites and general factor repertoires." EMBO J **26**(4): 944-954.

des Georges, A., Y. Hashem, A. Unbehaun, R. A. Grassucci, D. Taylor, C. U. Hellen, T. V. Pestova and J. Frank (2014). "Structure of the mammalian ribosomal pre-termination complex associated with eRF1.eRF3.GDPNP." Nucleic Acids Res **42**(5): 3409-3418.

Drygin, D., W. G. Rice and I. Grummt (2010). "The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer." Annu Rev Pharmacol Toxicol **50**: 131-156.

Edman, P. and G. Begg (1967). "A protein sequenator." Eur J Biochem **1**(1): 80-91.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Engel, C., T. Gubbey, S. Neyer, S. Sainsbury, C. Oberthuer, C. Baejen, C. Bernecky and P. Cramer (2017). "Structural Basis of RNA Polymerase I Transcription Initiation." Cell **169**(1): 120-131.e122.

Engel, C., J. Plitzko and P. Cramer (2016). "RNA polymerase I-Rrn3 complex at 4.8 Å resolution." Nat Commun **7**: 12129.

Engel, C., S. Sainsbury, A. C. Cheung, D. Kostrewa and P. Cramer (2013). "RNA polymerase I structure and transcription regulation." Nature **502**(7473): 650-655.

Engel, K. L., S. L. French, O. V. Viktorovskaya, A. L. Beyer and D. A. Schneider (2015). "Spt6 Is Essential for rRNA Synthesis by RNA Polymerase I." Mol Cell Biol **35**(13): 2321-2331.

Fath, S., P. Milkereit, G. Peyroche, M. Riva, C. Carles and H. Tschochner (2001). "Differential roles of phosphorylation in the formation of transcriptional active RNA polymerase I." Proc Natl Acad Sci U S A **98**(25): 14334-14339.

Fernandez-Leiro, R. and S. H. Scheres (2016). "Unravelling biological macromolecules with cryo-electron microscopy." Nature **537**(7620): 339-346.

Fernández-Tornero, C., B. Böttcher, M. Riva, C. Carles, U. Steuerwald, R. W. Ruigrok, A. Sentenac, C. W. Müller and G. Schoehn (2007). "Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III." Mol Cell **25**(6): 813-823.

Fernández-Tornero, C., M. Moreno-Morcillo, U. J. Rashid, N. M. Taylor, F. M. Ruiz, T. Gruene, P. Legrand, U. Steuerwald and C. W. Müller (2013). "Crystal structure of the 14-subunit RNA polymerase I." Nature **502**(7473): 644-649.

Friedrich, J. K., K. I. Panov, P. Cabart, J. Russell and J. C. Zomerdijk (2005). "TBP-TAF complex SL1 directs RNA polymerase I pre-initiation complex formation and stabilizes upstream binding factor at the rDNA promoter." J Biol Chem **280**(33): 29551-29558.

Frottin, F., A. Martinez, P. Peynot, S. Mitra, R. C. Holz, C. Giglione and T. Meinnel (2006). "The proteomics of N-terminal methionine cleavage." Mol Cell Proteomics **5**(12): 2336-2349.

Gallego, O., T. Specht, T. Brach, A. Kumar, A. C. Gavin and M. Kaksonen (2013). "Detection and characterization of protein interactions in vivo by a simple live-cell imaging method." PLoS One **8**(5): e62195.

Geiger, S. R., K. Lorenzen, A. Schreieck, P. Hanecker, D. Kostrewa, A. J. Heck and P. Cramer (2010). "RNA polymerase I contains a TFIIF-related DNA-binding subcomplex." Mol Cell **39**(4): 583-594.

Gerber, J., A. Reiter, R. Steinbauer, S. Jakob, C. D. Kuhn, P. Cramer, J. Griesenbeck, P. Milkereit and H. Tschochner (2008). "Site specific phosphorylation of yeast RNA polymerase I." Nucleic Acids Res **36**(3): 793-802.

Ghaemmaghami, S., W. K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea and J. S. Weissman (2003). "Global analysis of protein expression in yeast." Nature **425**(6959): 737-741.

Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler and A. S. Ponticelli (2004). "Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II." Mol Cell Biol **24**(24): 10975-10985.

Goetze, H., M. Wittner, S. Hamperl, M. Hondele, K. Merz, U. Stoeckl and J. Griesenbeck (2010). "Alternative chromatin structures of the 35S rRNA genes in Saccharomyces cerevisiae provide a molecular basis for the selective recruitment of RNA polymerases I and II." Mol Cell Biol **30**(8): 2028-2045.

Goodfellow, S. J. and J. C. Zomerdijk (2013). "Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes." Subcell Biochem **61**: 211-236.

Grohmann, D., A. Hirtreiter and F. Werner (2009). "RNAP subunits F/E (RPB4/7) are stably associated with archaeal RNA polymerase: using fluorescence anisotropy to monitor RNAP assembly in vitro." Biochem J **421**(3): 339-343.

Gross, C. A., C. Chan, A. Dombroski, T. Gruber, M. Sharp, J. Tupy and B. Young (1998). "The functional and regulatory roles of sigma factors in transcription." Cold Spring Harb Symp Quant Biol **63**: 141-155.

Gruber, T. M. and C. A. Gross (2003). "Multiple sigma subunits and the partitioning of bacterial transcription space." Annu Rev Microbiol **57**: 441-466.

Gu, W. and D. Reines (1995). "Variation in the size of nascent RNA cleavage products as a function of transcript length and elongation competence." J Biol Chem **270**(51): 30441-30447.

Hanada, K., C. Z. Song, K. Yamamoto, K. Yano, Y. Maeda, K. Yamaguchi and M. Muramatsu (1996). "RNA polymerase I associated factor 53 binds to the nucleolar transcription factor UBF and functions in specific rDNA transcription." EMBO J **15**(9): 2217-2226.

Hantsche, M. and P. Cramer (2017). "Conserved RNA polymerase II initiation complex structure." Curr Opin Struct Biol **47**: 17-22.

He, Y., J. Fang, D. J. Taatjes and E. Nogales (2013). "Structural visualization of key steps in human transcription initiation." Nature **495**(7442): 481-486.

Henderson, R. (2015). "Overview and future of single particle electron cryomicroscopy." Arch Biochem Biophys **581**: 19-24.

Hirata, A., B. J. Klein and K. S. Murakami (2008). "The X-ray crystal structure of RNA polymerase from Archaea." Nature **451**(7180): 851-854.

Hirata, A. and K. S. Murakami (2009). "Archaeal RNA polymerase." Curr Opin Struct Biol **19**(6): 724-731.

Hoffmann, N. A., A. J. Jakobi, M. Moreno-Morcillo, S. Glatt, J. Kosinski, W. J. Hagen, C. Sachse and C. W. Müller (2015). "Molecular structures of unbound and transcribing RNA polymerase III." Nature **528**(7581): 231-236.

Holstege, F. C., D. Tantin, M. Carey, P. C. van der Vliet and H. T. Timmers (1995). "The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA." EMBO J **14**(4): 810-819.

Holstege, F. C., P. C. van der Vliet and H. T. Timmers (1996). "Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH." EMBO J **15**(7): 1666-1677.

Hsin, J. P. and J. L. Manley (2012). "The RNA polymerase II CTD coordinates transcription and RNA processing." Genes Dev **26**(19): 2119-2137.

Huet, J., F. Wyers, J. M. Buhler, A. Sentenac and P. Fromageot (1976). "Association of RNase H activity with yeast RNA polymerase A." Nature **261**(5559): 431-433.

Jantzen, H. M., A. Admon, S. P. Bell and R. Tjian (1990). "Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins." Nature **344**(6269): 830-836.

Jeon, C., H. Yoon and K. Agarwal (1994). "The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II." Proc Natl Acad Sci U S A **91**(19): 9106-9110.

Jeronimo, C., P. Collin and F. Robert (2016). "The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain." J Mol Biol **428**(12): 2607-2622.

Kaplan, C. D. and R. D. Kornberg (2008). "A bridge to transcription by RNA polymerase." J Biol **7**(10): 39.

Keener, J., J. A. Dodd, D. Lalo and M. Nomura (1997). "Histones H3 and H4 are components of upstream activation factor required for the high-level transcription of yeast rDNA by RNA polymerase I." Proc Natl Acad Sci U S A **94**(25): 13458-13462.

Kettenberger, H., K. J. Armache and P. Cramer (2003). "Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage." Cell **114**(3): 347-357.

Keys, D. A., B. S. Lee, J. A. Dodd, T. T. Nguyen, L. Vu, E. Fantino, L. M. Burson, Y. Nogi and M. Nomura (1996). "Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex." Genes Dev **10**(7): 887-903.

Khaperskyy, D. A., M. L. Ammerman, R. C. Majovski and A. S. Ponticelli (2008). "Functions of Saccharomyces cerevisiae TFIIF during transcription start site utilization." Mol Cell Biol **28**(11): 3757-3766.

Knutson, B. A. and S. Hahn (2013). "TFIIB-related factors in RNA polymerase I transcription." Biochim Biophys Acta **1829**(3-4): 265-273.

Korkhin, Y., U. M. Unligil, O. Littlefield, P. J. Nelson, D. I. Stuart, P. B. Sigler, S. D. Bell and N. G. Abrescia (2009). "Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure." PLoS Biol **7**(5): e1000102.

Kostrewa, D., C. D. Kuhn, C. Engel and P. Cramer (2015). "An alternative RNA polymerase I structure reveals a dimer hinge." Acta Crystallogr D Biol Crystallogr **71**(Pt 9): 1850-1855.

Kuhn, C. D., S. R. Geiger, S. Baumli, M. Gartmann, J. Gerber, S. Jennebach, T. Mielke, H. Tschochner, R. Beckmann and P. Cramer (2007). "Functional architecture of RNA polymerase I." Cell **131**(7): 1260-1272.

Kulak, N. A., G. Pichler, I. Paron, N. Nagaraj and M. Mann (2014). "Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells." Nat Methods **11**(3): 319-324.

Kusnadi, E. P., K. M. Hannan, R. J. Hicks, R. D. Hannan, R. B. Pearson and J. Kang (2015). "Regulation of rDNA transcription in response to growth factors, nutrients and energy." Gene **556**(1): 27-34.

Kühlbrandt, W. (2014). "Biochemistry. The resolution revolution." Science **343**(6178): 1443-1444.

Laferté, A., E. Favry, A. Sentenac, M. Riva, C. Carles and S. Chédin (2006). "The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components." Genes Dev **20**(15): 2030-2040.

Lalo, D., J. S. Steffan, J. A. Dodd and M. Nomura (1996). "RRN11 encodes the third subunit of the complex containing Rrn6p and Rrn7p that is essential for the initiation of rDNA transcription by yeast RNA polymerase I." J Biol Chem **271**(35): 21062-21067.

Langer, D., J. Hain, P. Thuriaux and W. Zillig (1995). "Transcription in archaea: similarity to that in eucarya." Proc Natl Acad Sci U S A **92**(13): 5768-5772.

Laribee, R. N., A. Hosni-Ahmed, J. J. Workman and H. Chen (2015). "Ccr4-not regulates RNA polymerase I transcription and couples nutrient signaling to the control of ribosomal RNA biogenesis." PLoS Genet **11**(3): e1005113.

Laue, T. M., B. D. Shah, T. M. Ridgeway and S. L. Pelletier (1992). "Computer-aided interpretation of analytical sedimentation data for proteins." In: Harding SE, Rowe AJ, Horton JC, editors. Analytical ultracentrifugation in biochemistry and polymer science. Cambridge, UK: Royal Society of Chemistry: 90–125.

Learned, R. M., T. K. Learned, M. M. Haltiner and R. T. Tjian (1986). "Human rRNA transcription is modulated by the coordinate binding of two factors to an upstream control element." Cell **45**(6): 847-857.

Li, X., P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard and Y. Cheng (2013). "Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM." Nat Methods **10**(6): 584-590.

Lin, C. W., B. Moorefield, J. Payne, P. Aprikian, K. Mitomo and R. H. Reeder (1996). "A novel 66-kilodalton protein complexes with Rrn6, Rrn7, and TATA-binding protein to promote polymerase I transcription initiation in Saccharomyces cerevisiae." Mol Cell Biol **16**(11): 6436-6443.

Lyumkis, D., A. F. Brilot, D. L. Theobald and N. Grigorieff (2013). "Likelihood-based classification of cryo-EM images using FREALIGN." J Struct Biol **183**(3): 377-388.

Mathew, R. and D. Chatterji (2006). "The evolving story of the omega subunit of bacterial RNA polymerase." Trends Microbiol **14**(10): 450-455.

Mayer, C., H. Bierhoff and I. Grummt (2005). "The nucleolus as a stress sensor: JNK2 inactivates the transcription factor TIF-IA and down-regulates rRNA synthesis." Genes Dev **19**(8): 933-941.

Mayer, C., J. Zhao, X. Yuan and I. Grummt (2004). "mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability." Genes Dev **18**(4): 423-434.

McMullan, G., A. R. Faruqi, D. Clare and R. Henderson (2014). "Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy." Ultramicroscopy **147**: 156-163.

McStay, B., M. W. Frazier and R. H. Reeder (1991). "xUBF contains a novel dimerization domain essential for RNA polymerase I transcription." Genes Dev **5**(11): 1957-1968.

McStay, B. and I. Grummt (2008). "The epigenetics of rRNA genes: from molecular to chromosome biology." Annu Rev Cell Dev Biol **24**: 131-157.

Merrick, M. J. (1993). "In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)." Mol Microbiol **10**(5): 903-909.

Milkereit, P., P. Schultz and H. Tschochner (1997). "Resolution of RNA polymerase I into dimers and monomers and their function in transcription." Biol Chem **378**(12): 1433-1443.

Milkereit, P. and H. Tschochner (1998). "A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription." EMBO J **17**(13): 3692-3703.

Miller, G., K. I. Panov, J. K. Friedrich, L. Trinkle-Mulcahy, A. I. Lamond and J. C. Zomerdijk (2001). "hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters." EMBO J **20**(6): 1373-1382.

Miller, J. E. and J. C. Reese (2012). "Ccr4-Not complex: the control freak of eukaryotic cells." Crit Rev Biochem Mol Biol **47**(4): 315-333.

Miller, O. L. and B. R. Beatty (1969). "Visualization of nucleolar genes." Science **164**(3882): 955-957.

Minakhin, L., S. Bhagat, A. Brunning, E. A. Campbell, S. A. Darst, R. H. Ebright and K. Severinov (2001). "Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly." Proc Natl Acad Sci U S A **98**(3): 892-897.

Mindell, J. A. and N. Grigorieff (2003). "Accurate determination of local defocus and specimen tilt in electron microscopy." J Struct Biol **142**(3): 334-347.

Moorefield, B., E. A. Greene and R. H. Reeder (2000). "RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human." Proc Natl Acad Sci U S A **97**(9): 4724-4729.

Moreno-Morcillo, M., N. M. Taylor, T. Gruene, P. Legrand, U. J. Rashid, F. M. Ruiz, U. Steuerwald, C. W. Müller and C. Fernández-Tornero (2014). "Solving the RNA polymerase I structural puzzle." Acta Crystallogr D Biol Crystallogr **70**(Pt 10): 2570-2582.

Moss, T., F. Langlois, T. Gagnon-Kugler and V. Stefanovsky (2007). "A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis." Cell Mol Life Sci **64**(1): 29-49.

Mougey, E. B., M. O'Reilly, Y. Osheim, O. L. Miller, A. Beyer and B. Sollner-Webb (1993). "The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes." Genes Dev **7**(8): 1609-1619.

Neyer, S., M. Kunz, C. Geiss, M. Hantsche, V. V. Hodirnau, A. Seybert, C. Engel, M. P. Scheffer, P. Cramer and A. S. Frangakis (2016). "Structure of RNA polymerase I transcribing ribosomal DNA genes." Nature **540**(7634):607-610.

Nguyen, l. X., Y. Lee, L. Urbani, P. J. Utz, A. W. Hamburger, J. B. Sunwoo and B. S. Mitchell (2015). "Regulation of ribosomal RNA synthesis in T cells: requirement for GTP and Ebp1." Blood **125**(16): 2519-2529.

Nogales, E. and S. H. Scheres (2015). "Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity." Mol Cell **58**(4): 677-689.

Németh, A. and G. Längst (2011). "Genome organization in and around the nucleolus." Trends Genet **27**(4): 149-156.

O'Sullivan, A. C., G. J. Sullivan and B. McStay (2002). "UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat." Mol Cell Biol **22**(2): 657-668.

Okamoto, T., S. Yamamoto, Y. Watanabe, T. Ohta, F. Hanaoka, R. G. Roeder and Y. Ohkuma (1998). "Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIEbeta subunit." J Biol Chem **273**(31): 19866-19876.

Okuda, M., Y. Watanabe, H. Okamura, F. Hanaoka, Y. Ohkuma and Y. Nishimura (2000). "Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface." EMBO J **19**(6): 1346-1356.

Orlicky, S. M., P. T. Tran, M. H. Sayre and A. M. Edwards (2001). "Dissociable Rpb4-Rpb7 subassembly of rna polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation." J Biol Chem **276**(13): 10097-10102.

Osheim, Y. N., S. L. French, M. L. Sikes and A. L. Beyer (2009). "Electron microscope visualization of RNA transcription and processing in Saccharomyces cerevisiae by Miller chromatin spreading." Methods Mol Biol **464**: 55-69.

Panov, K. I., J. K. Friedrich, J. Russell and J. C. Zomerdijk (2006). "UBF activates RNA polymerase I transcription by stimulating promoter escape." EMBO J **25**(14): 3310-3322.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." J Comput Chem **25**(13): 1605-1612.

Peyroche, G., P. Milkereit, N. Bischler, H. Tschochner, P. Schultz, A. Sentenac, C. Carles and M. Riva (2000). "The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3." EMBO J **19**(20): 5473-5482.

(a) Pilsl, M., C. Crucifix, G. Papai, F. Krupp, R. Steinbauer, J. Griesenbeck, P. Milkereit, H. Tschochner and P. Schultz (2016). "Structure of the initiation-competent RNA polymerase I and its implication for transcription." Nat Commun **7**: 12126.

(b) Pilsl, M., P. E. Merkl, P. Milkereit, J. Griesenbeck and H. Tschochner (2016). "Analysis of S. cerevisiae RNA Polymerase I Transcription In Vitro." Methods Mol Biol **1455**: 99-108.

Plaschka, C., M. Hantsche, C. Dienemann, C. Burzinski, J. Plitzko and P. Cramer (2016). "Transcription initiation complex structures elucidate DNA opening." Nature **533**(7603): 353-358.

Plaschka, C., L. Larivière, L. Wenzeck, M. Seizl, M. Hemann, D. Tegunov, E. V. Petrotchenko, C. H. Borchers, W. Baumeister, F. Herzog, E. Villa and P. Cramer (2015). "Architecture of the RNA polymerase II-Mediator core initiation complex." Nature **518**(7539): 376-380.

Prescott, E. M., Y. N. Osheim, H. S. Jones, C. M. Alen, J. G. Roan, R. H. Reeder, A. L. Beyer and N. J. Proudfoot (2004). "Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p." Proc Natl Acad Sci U S A **101**(16): 6068-6073.

Price, D. H., A. E. Sluder and A. L. Greenleaf (1989). "Dynamic interaction between a Drosophila transcription factor and RNA polymerase II." Mol Cell Biol **9**(4): 1465-1475.

Ren, D., L. Lei and Z. F. Burton (1999). "A region within the RAP74 subunit of human transcription factor IIF is critical for initiation but dispensable for complex assembly." Mol Cell Biol **19**(11): 7377-7387.

Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann and B. Séraphin (1999). "A generic protein purification method for protein complex characterization and proteome exploration." Nat Biotechnol **17**(10): 1030-1032.

Rodthongkum, N., Y. Chen, S. Thayumanavan and R. W. Vachet (2010). "Selective enrichment and analysis of acidic peptides and proteins using polymeric reverse micelles and MALDI-MS." Anal Chem **82**(20): 8686-8691.

Rohou, A. and N. Grigorieff (2015). "CTFFIND4: Fast and accurate defocus estimation from electron micrographs." J Struct Biol **192**(2): 216-221.

Rosenthal, P. B. and R. Henderson (2003). "Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy." J Mol Biol **333**(4): 721-745.

Rubinstein, J. L. and M. A. Brubaker (2015). "Alignment of cryo-EM movies of individual particles by optimization of image translations." J Struct Biol **192**(2): 188-195.

Rudloff, U., D. Eberhard, L. Tora, H. Stunnenberg and I. Grummt (1994). "TBP-associated factors interact with DNA and govern species specificity of RNA polymerase I transcription." EMBO J **13**(11): 2611-2616.

Ruprich-Robert, G. and P. Thuriaux (2010). "Non-canonical DNA transcription enzymes and the conservation of two-barrel RNA polymerases." Nucleic Acids Res **38**(14): 4559-4569.

Russell, J. and J. C. Zomerdijk (2005). "RNA-polymerase-I-directed rDNA transcription, life and works." Trends Biochem Sci **30**(2): 87-96.

Russell, J. and J. C. Zomerdijk (2006). "The RNA polymerase I transcription machinery." Biochem Soc Symp(73): 203-216.

Russo, C. J. and L. A. Passmore (2014). "Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy." Science **346**(6215): 1377-1380.

Sanij, E. and R. D. Hannan (2009). "The role of UBF in regulating the structure and dynamics of transcriptionally active rDNA chromatin." Epigenetics **4**(6): 374-382.

Santoro, R., K. M. Schmitz, J. Sandoval and I. Grummt (2010). "Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans." EMBO Rep **11**(1): 52-58.

Scheres, S. H. (2010). "Classification of structural heterogeneity by maximum-likelihood methods." Methods Enzymol **482**: 295-320.

Scheres, S. H. (2012). "RELION: implementation of a Bayesian approach to cryo-EM structure determination." J Struct Biol **180**(3): 519-530.

Scheres, S. H. (2014). "Beam-induced motion correction for sub-megadalton cryo-EM particles." Elife **3**: e03665.

Scheres, S. H. (2016). "Processing of Structurally Heterogeneous Cryo-EM Data in RELION." Methods Enzymol **579**: 125-157.

Scheres, S. H., H. Gao, M. Valle, G. T. Herman, P. P. Eggermont, J. Frank and J. M. Carazo (2007). "Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization." Nat Methods **4**(1): 27-29.

Scheres, S. H., R. Núñez-Ramírez, C. O. Sorzano, J. M. Carazo and R. Marabini (2008). "Image processing for electron microscopy single-particle analysis using XMIPP." Nat Protoc **3**(6): 977-990.

Schnapp, A., G. Schnapp, B. Erny and I. Grummt (1993). "Function of the growth-regulated transcription initiation factor TIF-IA in initiation complex formation at the murine ribosomal gene promoter." Mol Cell Biol **13**(11): 6723-6732.

Schröder, R. R. (2015). "Advances in electron microscopy: A qualitative view of instrumentation development for macromolecular imaging and tomography." Arch Biochem Biophys **581**: 25-38.

Schuck, P. (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." Biophys J **78**(3): 1606-1619.

Sekine, S., S. Tagami and S. Yokoyama (2012). "Structural basis of transcription by bacterial and eukaryotic RNA polymerases." Curr Opin Struct Biol **22**(1): 110-118.

Shevchenko, A., M. Wilm, O. Vorm and M. Mann (1996). "Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels." Anal Chem **68**(5): 850-858.

Siddiqi, I. N., J. A. Dodd, L. Vu, K. Eliason, M. L. Oakes, J. Keener, R. Moore, M. K. Young and M. Nomura (2001). "Transcription of chromosomal rRNA genes by both RNA polymerase I and II in yeast uaf30 mutants lacking the 30 kDa subunit of transcription factor UAF." EMBO J **20**(16): 4512-4521.

Sigworth, F. J., P. C. Doerschuk, J. M. Carazo and S. H. Scheres (2010). "An introduction to maximum-likelihood methods in cryo-EM." Methods Enzymol **482**: 263-294.

Sosunov, V., E. Sosunova, A. Mustaev, I. Bass, V. Nikiforov and A. Goldfarb (2003). "Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase." EMBO J **22**(9): 2234-2244.

Stefanovsky, V., F. Langlois, T. Gagnon-Kugler, L. I. Rothblum and T. Moss (2006). "Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling." Mol Cell **21**(5): 629-639.

Stefanovsky, V. Y., G. Pelletier, D. P. Bazett-Jones, C. Crane-Robinson and T. Moss (2001). "DNA looping in the RNA polymerase I enhancesome is the result of non-cooperative in-phase bending by two UBF molecules." Nucleic Acids Res **29**(15): 3241-3247.

Steffan, J. S., D. A. Keys, J. A. Dodd and M. Nomura (1996). "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 Dev **10**(20): 2551-2563.

Steffan, J. S., D. A. Keys, L. Vu and M. Nomura (1998). "Interaction of TATA-binding protein with upstream activation factor is required for activated transcription of ribosomal DNA by RNA polymerase I in Saccharomyces cerevisiae in vivo." Mol Cell Biol **18**(7): 3752-3761.

Stepanchick, A., H. Zhi, A. H. Cavanaugh, K. Rothblum, D. A. Schneider and L. I. Rothblum (2013). "DNA binding by the ribosomal DNA transcription factor rrn3 is essential for ribosomal DNA transcription." J Biol Chem **288**(13): 9135-9144.

Tafur, L., Y. Sadian, N. A. Hoffmann, A. J. Jakobi, R. Wetzel, W. J. Hagen, C. Sachse and C. W. Müller (2016). "Molecular Structures of Transcribing RNA Polymerase I." Mol Cell **64**(6):1135:1143.

Tan, S., T. Aso, R. C. Conaway and J. W. Conaway (1994). "Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II." J Biol Chem **269**(41): 25684-25691.

Tanaka, A., T. Watanabe, Y. Iida, F. Hanaoka and Y. Ohkuma (2009). "Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation." Genes Cells **14**(3): 395-405.

Tang, G., L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees and S. J. Ludtke (2007). "EMAN2: an extensible image processing suite for electron microscopy." J Struct Biol **157**(1): 38-46.

Torreira, E., J. A. Louro, I. Pazos, N. González-Polo, D. Gil-Carton, A. G. Duran, S. Tosi, O. Gallego, O. Calvo and C. Fernández-Tornero (2017). "The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription." Elife **6**.

Tseng, H., W. Chou, J. Wang, X. Zhang, S. Zhang and R. M. Schultz (2008). "Mouse ribosomal RNA genes contain multiple differentially regulated variants." PLoS One **3**(3): e1843.

Vannini, A. and P. Cramer (2012). "Conservation between the RNA polymerase I, II, and III transcription initiation machineries." Mol Cell **45**(4): 439-446.

Vu, L., I. Siddiqi, B. S. Lee, C. A. Josaitis and M. Nomura (1999). "RNA polymerase switch in transcription of yeast rDNA: role of transcription factor UAF (upstream activation factor) in silencing rDNA transcription by RNA polymerase II." Proc Natl Acad Sci U S A **96**(8): 4390-4395.

Walmacq, C., M. L. Kireeva, J. Irvin, Y. Nedialkov, L. Lubkowska, F. Malagon, J. N. Strathern and M. Kashlev (2009). "Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II." J Biol Chem **284**(29): 19601-19612.

Warner, J. R. (1999). "The economics of ribosome biosynthesis in yeast." Trends Biochem Sci **24**(11): 437-440.

Weinzierl, R. O. (2011). "The Bridge Helix of RNA polymerase acts as a central nanomechanical switchboard for coordinating catalysis and substrate movement." Archaea **2011**: 608385.

Werner, F. and D. Grohmann (2011). "Evolution of multisubunit RNA polymerases in the three domains of life." Nat Rev Microbiol **9**(2): 85-98.

Yamamoto, K., M. Yamamoto, K. Hanada, Y. Nogi, T. Matsuyama and M. Muramatsu (2004). "Multiple protein-protein interactions by RNA polymerase I-associated factor PAF49 and role of PAF49 in rRNA transcription." Mol Cell Biol **24**(14): 6338-6349.

Yan, Q., R. J. Moreland, J. W. Conaway and R. C. Conaway (1999). "Dual roles for transcription factor IIF in promoter escape by RNA polymerase II." J Biol Chem **274**(50): 35668-35675.

Yuan, X., J. Zhao, H. Zentgraf, U. Hoffmann-Rohrer and I. Grummt (2002). "Multiple interactions between RNA polymerase I, TIF-IA and TAF(I) subunits regulate preinitiation complex assembly at the ribosomal gene promoter." EMBO Rep **3**(11): 1082-1087.

Zhang, G., E. A. Campbell, L. Minakhin, C. Richter, K. Severinov and S. A. Darst (1999). "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution." Cell **98**(6): 811-824.

CONTRIBUTION TO SCIENTIFIC ARTICLES AS PhD STUDENT

ARTICLE RELATED TO THE DOCTORAL THESIS

Torreira, E., J. A. Louro, I. Pazos, N. González-Polo, D. Gil-Carton, A. G. Duran, S. Tosi, O. Gallego, O. Calvo and C. Fernández-Tornero (2017). "The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription." Elife **6**: e20832.
(attached in the next page)

ADDITTIONAL ARTICLE

Garavís, M., N. González-Polo, P. Allepuz-Fuster, J. A. Louro, C. Fernández-Tornero, O. Calvo (2017). "Sub1 contacts the RNA polymerase II stalk to modulate mRNA synthesis." Nucleic Acids Res **45**(5): 2458-2471.