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

submitted by
Matthias Kopano Vorländer, MSc
born in Haan, Germany
Oral examination: 15th of May 2019

Structural studies of RNA Polymerase III transcription

Referees:
Prof. Dr. Carsten Sachse
Prof. Dr. Irmgard Sinning

ABSTRACT

RNA Polymerase III (Pol III) produces small, non-coding RNAs with fundamental functions in the eukaryotic cell, including translation, splicing and protein sorting. While structures of unbound and transcribing Pol III have been solved and provided valuable mechanistic insights into Pol III transcription, snapshots of molecular interactions that underlie Pol III activation and repression are lacking. In this thesis I address these questions with structural studies of the *Sarrachomyces cerevisiae* Pol III transcription apparatus. I present high-resolution cryo-EM reconstructions of Pol III bound to its principal transcription initiation factor TFIIIB that were used to build atomic models. The complex was observed in different functional states, including two early intermediates in which the DNA duplex is closed, an open DNA complex, and an initially transcribing complex with RNA in the active site. The structures reveal an extremely tight, multivalent interaction between TFIIIB and promoter DNA, and explain how TFIIIB recruits Pol III. Together, TFIIIB and Pol III subunit C37 activate the intrinsic transcription factor-like function of the Pol III-specific heterotrimer to initiate the melting of double-stranded DNA, in a mechanism similar to that of the Pol II system.

I further present a high resolution structure of Pol III bound to the negative regulator Maf1, that explains how Maf1 achieves transcription repression by preventing interaction with TFIIIB. Maf1 occupies a position on Pol III that overlaps with the binding site of promoter DNA and TFIIIB. Furthermore, by mimicking the shape and electrostatic charge of a double-stranded DNA backbone, Maf1 further sequesters a mobile domain of Pol III subunit C34, which seals off the active site cleft and makes it inaccessible to bind DNA.

Lastly, I describe a recombinant expression system for the six-subunit, 520 kDa transcription factor TFIIIC and subcomplexes thereof. Negative stain electron microscopy of a complex between the τA module of TFIIIC and TFIIIB provide the first molecular insights into how TFIIIC recruits TFIIIB and positions it upstream of the transcription start site. Biochemical experiments further show that the τA module is displaced after or concomitant with Pol III recruitment, establishing it as an assembly factor rather than a *bona fide* transcription factor.

ZUSAMMENFASSUNG

RNA Polymerase III (Pol III) produziert kleine, nicht-kodierende RNAs, die fundamentale Aufgaben in der eukaryotischen Zelle ausüben, inklusive Translation, Splicing und Proteinsortierung. Strukturen von ungebundener und transkribierender Pol III wurden gelöst und lieferten wertvolle mechanistische Einsichten in die Pol III Transkription. Schnappschüsse der molekularen Interaktionen, die der Aktivierung und Repression von Pol III zugrunde liegen, fehlen hingegen. In dieser Arbeit gehe ich diesen offenen Fragen durch strukturelle Untersuchungen des *Sacchaaromyces cerevisiae* Pol III Transpkriptionsapparates nach. Ich zeige hochaufgelöste cryo-EM Rekonstruktionen von Pol III, gebunden an den wichtigsten Transkriptionsinitiationsfaktor TFIIIB, und habe atomare Modelle des TFIIIB-Pol III Komplexes erstellt. Der Komplex wurde in verschiedenen funktionellen Zuständen beobachtet. Darunter sind zwei frühe Intermediate, in denen die DNA Duplex geschlossen ist, ein offener DNA-Komplex und ein initial transkribierender Komplex mit RNA im aktiven Zentrum. Die Strukturen legen eine sehr enge, multivalente Interaktion zwischen TFIIIB und Promotor-DNA offen und erklären, wie Pol III von TFIIIB rekrutiert wird. TFIIIB und die Pol III-Untereinheit C37 aktivieren zusammen die intrinsische transkriptionsfaktorartige Aktivität des Pol III-spezifischen Heterotrimers, um die doppelsträngige DNA aufzuschmelzen. Dieser Mechanismus weist Ähnlichkeiten zum Pol II System auf.

Weiterhin präsentiere ich Strukturen von Pol III, gebunden an den negativen Regulator Maf1. Diese zeigen, wie Maf1 Transkriptionsrepression bewirkt, indem es die Interaktion mit TFIIIB verhindert. Maf1 bindet Pol III an einer Stelle die mit der Position von TFIIIB und Pomotor-DNA überlappt. Außerdem ähnelt Maf1 in Form und elektrostatischer Ladung dem Rückgrad von doppelsträngiger DNA, und bindet dadurch eine mobile Domäne der Pol III Untereinheit C34, wodurch das aktive Zentrum für DNA unzugänglich wird.

Schlussendlich beschreibe ich ein rekombinantes Expressionssystem für den Transktiptionsfaktor TFIIIC, der aus sechs Untereinheiten besteht und ein Molekulargewicht von 520 kDa aufweist. Eine *negative stain* elektronenmikroskopische Rekonstruktion des τA-Moduls von TFIIIC in einem Komplex mit TFIIIB liefern die ersten molekularen Einsichten in die Frage, wie TFIIIB von TFIIIC rekrutiert und TFIIIB oberhalb der Transkriptionsstartstelle positioniert wird. Biochemische Experimente zeigen außerdem, dass das τA-Modul von Pol III verdrängt wird, wenn Pol III an TFIIIB bindet. Dieses Ergebnis zeigt, dass τA eher als Assemblierungsfaktor denn als *bona fide* Transkriptionsfaktor fungiert.

PUBLICATIONS

1. Hoffmann, N. a., Jakobi, A. J., Vorländer, M. K., Sachse, C. & Müller, C. W. Transcribing RNA polymerase III observed by electron cryomicroscopy. *FEBS Journal,* 2811–2819 (2016).
2. Khatter, H., Vorländer, M. K. & Müller, C. W. RNA polymerase I and III: similar yet unique. *Current Opinion in Structural Biology* **47,** 88–94 (2017).
3. Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553,** 295–300 (2018).

Standing on the shoulders of Giants

— Bernard of Chartres

ACKNOWLEDGMENTS

The quote above expresses my deep gratitude to all those who have made this dissertation possible. It refers to the decades of scientists that have provided the technological and scientific foundation of this thesis, by building the microscopes and writing the software that produces these astonishing high resolution images, which allowed me to capture a snapshot of transcription. Equally important to interpret the data was the biochemical knowledge that was available, painstakingly gathered over many years by many laboratories providing a fine functional dissection of the Pol III transcription apparatus.

I am equally grateful to my supervisor, Dr. Christoph Müller, who has given me the opportunity to work on this exciting project and gave me a lot of freedom to do so. I appreciate his patience when data had to be collected, models built, figures made and a manuscript written at a time where our friendly competitors where a small step ahead of us.

There are many more people without who this work would not have been possible. First, Niklas Hoffmann helped to get me started with this project and taught me how to prepare samples and analyse data. Rene Wetzel, Helga Groetsch, Yashar Sardian, Lucas Tafur and Heena Khatter all helped to get the proteins required for this. I thank Florence Baudin for her invaluable help with all ’hot’ biochemical experiments, and all other members of the Müller group for a great environment in the lab. I am very luck to have worked with Wim Hagen and Felix Weiss, who run one of the best electron microscopy facilities in the world. Lastly, Anna Jungblut has been an incredible help for many experiments and is responsible for obtaining many of the results presented in the last chapter of this thesis due to her diligence and enthusiasm.

I am exceedingly thankful to the friends that I have made during the last 4 years in Heidelberg, who have laughed and played and danced and travelled and shared a home with me and have made this an unforgettable time, which I will be looking back to sentimentally.

Lastly, I want to say thank you to my parents. You have always shown me unconditional trust and love, you are infinitely curious about the scientific world, and you truly are the giants upon whose shoulders I stand.

CONTENTS

I INTRODUCTION

1 POL III GENES AND THE POL III TRANSCRIPTION APPARATUS: A MOLECULAR PERSPECTIVE 3
- 1.1 Pol III is specialized in high-level transcription of short RNAs 3
- 1.2 Pol III genes 4
- 1.3 The Pol III transcription apparatus 5
  - 1.3.1 The Pol III enzyme 5
  - 1.3.2 TFIIIB 7
  - 1.3.3 TFIIIC 9
  - 1.3.4 Pre-initiation complex formation and promoter opening 13
  - 1.3.5 Transcription repression by Maf1 14
- 1.4 Pol III promoter types 15
  - 1.4.1 Type I promoters - 5S rRNA 15
  - 1.4.2 Type II promoters - tRNAs 16
  - 1.4.3 Type III promoters 16
- 1.5 Pol III transcription in the context of the nucleus 17
  - 1.5.1 Gene occupancy by the Pol III machinery *in vivo* 17
  - 1.5.2 Extra TFIIIC sites and tDNAs as insulators 18
  - 1.5.3 Subnuclear localization of Pol III transcription 19
- 1.6 Pol III transcription and disease 19
  - 1.6.1 Pol III and cancer 19
  - 1.6.2 Pol III-associated developmental and neurological diseases 20
  - 1.6.3 Pol III and links to immunity 21
- 1.7 Using electron microscopy to study transcription complexes 22
  - 1.7.1 The cryo-EM structure determination process 23
  - 1.7.2 Typical workflow of a single particle cryo-EM structure determination project 24
- 1.8 Aims and motivation of this thesis 27

II RESULTS

2 STRUCTURAL MECHANISM OF PROMOTER OPENING IN THE RNA POLYMERASE III PRE-INITIATION COMPLEX 31
- 2.1 Reconstitution of the TFIIIB-Pol III complex 31
  - 2.1.1 Expression and purification of TFIIIB subunits 31
  - 2.1.2 Reconstitution of the TFIIIB-Pol III complex 33
- 2.2 Cryo-EM of the Pol III-TFIIIB complex 35
  - 2.2.1 Optimization of cryo-EM sample preparation of the TFIIIB-Pol III complex 35
  - 2.2.2 Data processing and particle classification 37
  - 2.2.3 High resolution cryo-EM maps of the TFIIIB-Pol III structure in different functional states 37
  - 2.2.4 Architecture of the Pol III open complex / initially transcribing complex 43

2.2.5 DNA binding mode of TFIIIB explains long residence time on chromatin . . . 46
2.2.6 TFIIIB coordinates promoter opening . . . 48
2.2.7 The transcription bubble is stabilized by Pol III, but coordinated by TFIIIB . . . 49
2.2.8 Structures of the closed DNA complexes . . . 51
2.2.9 Mechanistic model of promoter opening . . . 52
3 MOLECULAR BASIS OF POL III REPRESSION BY MAF1 55
3.1 Reconstitution of the Maf1-Pol III complex and initial EM reconstructions . 55
3.2 Crystallization of Maf1 . . . 56
3.3 Optimization of the Maf1-Pol III sample for high-resolution cryo-EM 59
3.4 Particle classification of the Maf1 dataset . . . 61
3.5 Structure of the Maf1-Pol III complex . . . 62
3.6 Mechanism of repression by Maf1 . . . 65
4 ASSEMBLY OF TFIIIB BY TFIIIC 67
4.1 Recombinant expression of TFIIIC . . . 67
4.2 Efforts to reconstitute the TFIIIC-TFIIIB-Pol III complex . . . 69
4.2.1 Efforts to reconstitute the TFIIIC-TFIIIB-TFIIIA-Pol III complex on a type I promoter . . . 70
4.2.2 Reconstitution of a TFIIIB-TFIIIC complex . . . 71
4.3 Structural analysis of the τA subcomplex . . . 71
4.3.1 Towards a high resolution cryo-EM structure of the τA-TFIIIB complex 74
4.3.2 Comparison of the τA-TFIIIB map with negative stain maps of TFIIIC and TFIIIC-TFIIIB . . . 76
4.3.3 τA is displaced from the Pol III-TFIIIB complex during promoter opening . . . 77

III DISCUSSION

5 DISCUSSION 85
5.1 The architecture of the Pol III PIC is adapted to high-frequency initiation . 85
5.1.1 Comparison of the Pol III PIC architecture with Pol II and Pol I . . . 85
5.2 Maf1 prevents Pol III transcription initiation by steric hindrance . . . 87
5.3 Towards a cryo-EM structure of TFIIIC . . . 89
5.4 TFIIIC functions as an assembly factor . . . 91

IV MATERIAL AND METHODS

6 MATERIAL AND METHODS 95
6.1 Cloning of recombinant TFIIIC . . . 95
6.2 Protein expression and purification . . . 95
6.2.1 Pol III . . . 95
6.2.2 Brf1-TBP . . . 96
6.2.3 Bdp1 . . . 96
6.2.4 Maf1 . . . 97
6.2.5 TFIIIA . . . 97
6.2.6 Recombinant TFIIIC . . . 97
6.2.7 Recombinant τA . . . 98
6.3 Structure determination of the Pol III-TFIIIB complex . . . 98

6.3.1 DNA oligo-nucleotides . . . 98
6.3.2 Assembly of pre-initiation complexes . . . 99
6.3.3 Electron microscopy and data processing . . . 99
6.3.4 Model building . . . 100
6.3.5 Local amplitude scaling (LocScale) . . . 102
6.3.6 Model refinement and validation . . . 102
6.3.7 Modelling of the open clamp, closed complex structure . . . 102
6.3.8 RNA extension assay . . . 103
6.3.9 Data availability . . . 103
6.4 Structure determination of the Maf1-Pol III complex . . . 103
6.4.1 Cryo-EM sample preparation . . . 103
6.4.2 Cryo-EM imaging and pre-processing . . . 104
6.4.3 Maf1-Pol III particle classification . . . 104
6.4.4 Model building . . . 104
6.4.5 Maf1-C34 pull-down . . . 106
6.5 TFIIIC . . . 106
6.5.1 Reconstitution of the τA -TFIIIB complex . . . 106
6.5.2 Negative stain EM structure data collection and processing . . . 106
6.5.3 Transcsription assay . . . 109

V APPENDIX

BIBLIOGRAPHY 113

LIST OF FIGURES

Figure 1.1 Structure of Pol III . . . 5
Figure 1.2 Comparison of Pol I, II and III, and the Pol II PIC . . . 7
Figure 1.3 TFIIIB subunits and domains . . . 8
Figure 1.4 Architecture of TFIIIC and interaction with TFIIIB . . . 11
Figure 2.1 Optimization of Brf1-TBP expression and purification . . . 32
Figure 2.2 Reconstitution of the Pol III-TFIIIB complex . . . 34
Figure 2.3 EM sample optimization of the Pol III PIC . . . 36
Figure 2.4 Classification of particles from the ITC datset . . . 38
Figure 2.5 Classification of particles from the CC datset . . . 39
Figure 2.6 Cryo-EM densities of the Pol III ITC, OC, CC1 and CC2 maps . . . 40
Figure 2.7 Cryo-EM maps of the Pol III PIC colored by local resolution . . . 42
Figure 2.8 Overview of the open complex Pol III-PIC . . . 43
Figure 2.9 Structures of Brf1 and Bdp1 in the PIC . . . 44
Figure 2.10 RNA extension assay . . . 45
Figure 2.11 TFIIIB forms a positively charged ring around upstream promoter DNA . . . 47
Figure 2.12 Allosteric changes in PIC . . . 49
Figure 2.13 DNA binding around the transcription bubble . . . 50
Figure 2.14 TFIIIB interactions with C34 and C37 in the PIC . . . 51
Figure 2.15 Comparison between Pol III closed complex and open complex structures . . . 52
Figure 2.16 Model of promoter opening in the Pol III PIC . . . 54
Figure 3.1 Full-length Maf1 aggregates under low-salt conditions . . . 55
Figure 3.2 Initital cryo-EM reconstruction of Maf1-Pol III . . . 56
Figure 3.3 Crystallization of Maf1 . . . 58
Figure 3.4 Comparison of crosslinked and non-crosslinked MAf1-Pol III reconstructions . . . 60
Figure 3.5 Particle classification of the Maf1-Pol III dataset . . . 61
Figure 3.6 Multibody refinement of Maf1-Pol III . . . 62
Figure 3.7 Structure of the Maf1-Pol III complex . . . 63
Figure 3.8 Close-up of the Maf1-Pol III interaction . . . 64
Figure 3.9 Side-by-side comparison of the Pol III conformation . . . 64
Figure 3.10 Maf1 mimics promoter DNA to bind C34 . . . 65
Figure 3.11 Maf1 blocks the TFIIIB binding site . . . 66
Figure 4.1 Insect cell expression of TFIIIC . . . 68
Figure 4.2 Recombinant TFIIIC is active . . . 69
Figure 4.3 Negative stain data analysis of a sample of TFIIIC, TFIIIB and Pol III 70
Figure 4.4 *In vitro* transcription of various Pol III promoter sequences . . . 71
Figure 4.5 τA stimulates basic levels of transcription and stoichiometrically binds TFIIIB . . . 72
Figure 4.6 Negative stain analysis of the τA-TFIIIB complex and τA only . . . 73
Figure 4.7 Negative stain density fit of τA and TFIIIB . . . 74

Figure 4.8 2D classes of cryo datasets collected of τA and TFIIIB 76
Figure 4.9 Comparison of negative stain reconstructions of TFIIIC and TFIIIC-TFIIIB 77
Figure 4.10 2D classes of cryo datasets collected of τA and TFIIIB 78
Figure 4.11 Model of the TFIIIC-PIC 79
Figure 4.12 *In vitro* transcription of his-tRNA promoter mutants 80
Figure 4.13 SEC and negative stain analysis of Pol III, Bdp1Δ, Brf1-TBPΔ and mutated his promoter 81
Figure 5.1 Comparison of Pol I, II and III PICs 86
Figure 5.2 The Bdp1 SANT domain is located at a similar position as TFIIA . 88
Figure 5.3 The Maf1 acidic patch is conserved 89
Figure 5.4 Positions of engineered mutations in Maf1 and Pol III 90
Figure 6.1 Data collection and refinement statistics of the Maf1-Pol III complex 105
Figure 6.2 Data collection and processing for the τA-TFIIIB negative stain structure 108

LIST OF TABLES

Table 2.1 Cryo-EM datasets acquired for the Pol III-TFIIIB complex 37
Table 4.1 Summary of reconstitution experiments with mutant and wildtype TFIIIB and promoter DNA 81
Table 6.1 Expression construct for recombinant TFIIIC production in insect cells 95
Table 6.2 Data collection and model refinement statistics for the Pol III-TFIIIB complex 101

ABBREVIATIONS

RNP	ribonucleoprotein
TF	transcription factor
ICR	internal control region
TEs	transposable elements
SINEs	short interspersed elements
rDNA	ribosomal DNA

ZF	Zinc-finger
TSS	transcription start site
PIC	pre-initiation complex
RNAP	DNA-dependent RNA Polymerase
TBP	TATA-binding protein
Brf1	B-related factor
TPR	tetratricopeptide repeat
HMG	high-mobility group
WH	winged-helix
eWH	extended winged-helix
ChIP	chromatin immunoprecipitation
ETC	extra TFIIIC
SEC	size exclusion chromatography
KOAc	potassium acetate
Bdp1	B double prime
nt	nucleotides
bp	base pairs
OC	open DNA complex
CC	closed DNA complex
ITC	initially transcribing complex

Part I

INTRODUCTION

1

POL III GENES AND THE POL III TRANSCRIPTION APPARATUS: A MOLECULAR PERSPECTIVE

Transcription is the process in which genetic information is read from DNA, producing different classes of functional RNA molecules. It is a key process that allows cellular plasticity, i.e. response to external stimuli and cellular differentiation are brought about by underlying changes in the transcriptome. While prokaryotes use one DNA-dependent RNA Polymerase (RNAP) to produce all transcripts, eukaryotes have evolved three (plants: five) RNAPs with different transcriptional profiles. RNA Polymerase I (Pol I) produces a single long ribosomal rRNA precursor, RNA Polymerase II (Pol II) transcribes mRNAs and short, regulatory RNAs, and RNA Polymerases III (Pol III) produces short, untranslated RNAs with extensive tertiary structure. In this introduction I will try to illuminate the role of Pol III by describing its transcriptome, the cellular machinery that underlies Pol III function, the way this machinery shapes chromatin structure, and how Pol III activity is regulated, including examples of diseases associated with miss-regulated Pol III transcription.

1.1 POL III IS SPECIALIZED IN HIGH-LEVEL TRANSCRIPTION OF SHORT RNAS

Pol III has a unique transcriptional profile among the eukaryotic RNAPs because it is specialized in high level transcription of very short genes, rarely exceeding 200 nucleotides (nt) (see section 1.2). The highly transcribed nature of Pol III genes becomes obvious when the density of Pol II and Pol III occupancies on their target genes are compared: The average Pol II density on yeast genes was estimated to be around one Pol II enzyme every 12 000 base pairs (bp) [156], whereas one Pol III enzyme is estimated to be found every 44 bp for the case of the 5S rRNA gene [60].

To achieve high levels of transcription on short genes, the Pol III transcription apparatus has evolved adaptations at all steps of the transcription cycle: initiation, elongation and termination. Initiation involves a sequence of events: assembly of general transcription factor (TF)s on Pol III genes, recruitment of Pol III, promoter opening and initial synthesis of RNA, and promoter escape. The initial synthesis of RNA is referred to as "abortive transcription", because often only short oligonucleotides are produced since Pol III fails to escape from the promoter. Once the nascent transcript reaches a critical size, the connection to TFIIIB is broken and Pol III enters the elongation phase. Finally, transcription has to be terminated by releasing the nascent transcript from the active site.

To perform rapid cycles of initiation, elongation and termination on the same genes, Pol III employs a feature termed facilitated recycling (reviewed in [5]) that is unique among eukaryotic RNAPs. This process has been shown to increase transcriptional output significantly, as the rate of transcription in subsequent cycles increases at least 5-fold after the initial round of transcription *in vitro*. To achieve this, Pol III (as Pol I, but in contrast to Pol II) does not require an ATP-dependent helicase for promoter opening. Pol III also does not accumulate post-transcriptional modifications (i.e. phosphorylation of the Pol II C-terminal domain) that mark the transitions into different phases of the transcription cycle. Moreover, subunits homologous to general Pol II transcription factors are stably integrated in the Pol III enzyme (see section 1.3) and therefore do not have to be pre-assembled on promoters prior to transcription initiation. Facilitated recycling also benefits from an efficient termination mechanism. While Pol I and Pol II require additional termination factors, Pol III terminates autonomously when it encounters a stretch of 5-7 thymines in the non-template strand [35]. Lastly, the principal transcription factor of Pol III—TFIIIB—remains bound to its promoters after release from Pol III. Consequently, 99% of Pol III initiation events are re-initiation on pre-assembled TFIIIB, whereas Pol II transcription factors typically initiate single rounds of transcription [5].

1.2 POL III GENES

Pol III produces untranslated RNAs, that often have catalytic functions as components of ancient ribonucleoprotein (RNP) complexes in the cell. While the most prominent class of Pol III transcripts, the tRNAs and the 5S rRNA, are essential components of the translation apparatus, other transcripts perform equally important cellular tasks. The U6 RNA is the most highly conserved part of the spliceosome [18] and forms its active site. The 7SL RNA forms the core of the signal recognition particle [48], which orchestrates co-translational import into the endoplasmatic reticulum, and RNAse P and the mitochondrial RNase P are required for tRNA and rRNA processing as well as mitochondrial DNA replication [72].

Interestingly, the different classes of transcripts are regulated by distinct promoter systems, termed type I (solely used by the 5S rRNA), type II (tRNA genes), and, for the most part an metazoan adaptation, type III (U6, 7SK, RNAse P, MRP, and others). The architecture of these promoters is discussed in section 1.4.

In addition to these long-known[1], highly abundant and stable transcripts—referred to as 'housekeeping genes'—recent technological advancements have allowed the discovery of more transient transcripts which exert regulatory functions, often influencing Pol II transcription. The complexity of these processes is increasing with organismal complexity, and many new functions of Pol III transcripts are discovered in animals that are absent in yeast and show tissue-specific expression [155].

For example, Pol III plays a fundamental role in the biology of transposable elements (TEs). TEs are a diverse class of mobile genetic elements, which make up an astonishing percentage of animal and plant genomes due to their self replicating nature. In humans, TEs constitute at least 45% of all DNA [123] and in maize this number increases to 90% [89]. TEs play a complex role in genome biology, because their insertion can disrupt genes or regulatory sequences and thereby cause disease[2]. On the other hand, they can act as cis-regulatory factors controlling gene expression [119] and are believed to be a driving force in genome evolution [180]. The most abundant class of TEs in humans [3], called short interspersed elements (SINEs), are all derived from Pol III genes and rely on transcription by Pol III due to their internal promoter elements [119]. A second class of TEs, the virus-like long terminal repeat retrotransposons, although not transcribed by Pol III, integrate into the genome close to Pol III genes in yeast, mediated by a direct interaction of the integrase with the Pol III subunits AC40 [17] and C53 [29].

---

[1] Extensively researched from 1970-1990

[2] 96 single-gene diseases caused by transposon insertion have been identified [74]

[3] The 7sL-derived Alu element alone is present in 1 100 000 copies in the human genome [74]

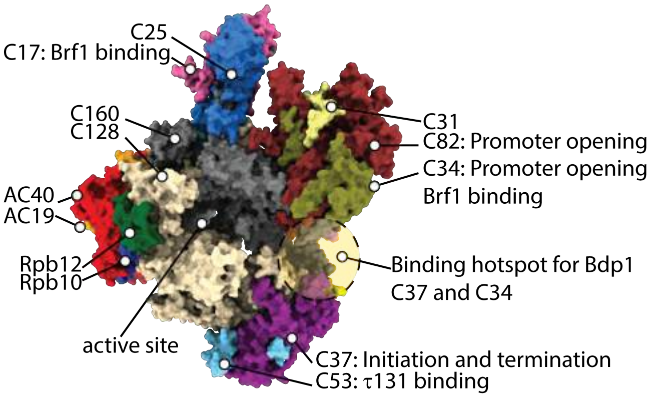


Figure 1.1: Structure of Pol III rendered based on PDB 5fja. Subunits and their interaction with transcription factors are labelled according to information presented in section 1.3.4

Finally, in the light of the recent wide-spread use of the CRISPR-Cas9 system, it should be mentioned that the guide RNAs used to target specific genes are transcribed under control of a Pol III promoter, and this has sparked a new interest in optimized Pol III promoters and terminators [46, 140, 184].

1.3 THE POL III TRANSCRIPTION APPARATUS

1.3.1 *The Pol III enzyme*

Pol III was initially identified as one of three RNA polymerase activities in fractionated sea urchin extracts [167]. Exploiting the different sensitivities of the three RNA polymerases to the toxin α-amanitin, it was soon established that Pol III transcribes tRNAs and 5S rRNA [136].

Pol III contains 17 subunits. Sequence homology to bacterial RNA Polymerase and Pol I and Pol II, and an extensive yeast two-hybrid screen [59] have established that the subunits are arranged in a 10-subunit core that is decorated by three peripheral subcomplexes; the C17-C25 stalk, the C37-C53 heterodimer, and the C31-C34-C82 heterotrimer. A high-resolution cryo-EM structure eventually provided an atomic model of the 17-subunit Pol III [82](Figure 1.1).

The core is the most highly conserved part. While the two biggest subunits which constitute the active site are unique for Pol I, Pol II and Pol III, they are highly conserved and homologous to the bacterial RNAP subunits β′ and β. Five subunits (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12) are shared between Pol I, Pol II and Pol III. They are located around the two biggest subunits [82]. Pol I and Pol III share two additional subunits (AC40 and AC19); both of these subunits are homologous to the bacterial RNAP α-subunit.

The core has a characteristic horseshoe-like shape which encloses a central cleft that harbors the DNA, and has two channels, one for the substrate NTPs and the other for the RNA product [207, 218]. Crystallographic and functional studies of yeast Pol II elucidated how a number of conserved, flexible elements in the active site maintain a transcription bubble with separated DNA strands, catalyze the addition of nucleotides, facilitate translocation along the template, stabilize the DNA:RNA hybrid and finally allow the DNA strands to re-anneal (reviewed in [19, 20]). The active site is highly conserved between Pol I, Pol II and Pol III, and the basic mechanism of transcription can thus be inferred from studies of the Pol II system [190].

Two additional domains, called jaw and clamp, stabilize the DNA at the downstream end and allow opening and closing of the active site cleft. In Pol III, the clamp head is enlarged compared to Pol I and Pol II, and this coincides with a narrow active site cleft.

Lastly, the core subunit C11 is a hybrid protein. The C11 N-terminal domain is tightly anchored to the Pol III core and has homology to the Pol II subunit Rpb9, but the C-terminal moiety is mobile and located outside of the Pol III core in available structures [82]. The C-terminal domain has homology with the Pol II transcription factor TFIIS, and is required for RNA cleavage and transcription termination [165]. The Pol I subunit A12 has a strikingly similar chimeric architecture and function as C11 [55, 173].

The peripheral stalk subcomplex is present in all eukaryotic and archael RNAPs, and in all cases interacts with initiation factors [201]. The Pol III C17-C25 stalk subunits are anchored to the core via extensions of the largest subunit C160 and a Pol III-specific helix of C25 that contacts the clamp domain [82]. The C17 subunit was further shown to interact with the TFIIIB subunit Brf1 (see section 1.3.4).

The C37-C53 heterodimer is anchored to the lobe of Pol III, opposite of the stalk. C53-C37 dimerize via a triple β-barrel fold, and similar dimarization domains are found in the Pol I A49-A34.5 hetrodimer and the Pol II general transcription factor TFIIF [55, 61, 82, 201]. However the Pol III heterodimer lacks the winged-helix (WH)-domains present in TFIIF and A49-A34.5.

The Pol III heterodimer has a dual function and is required for both transcription initiation and termination, and a stretch of residues in C37 (amino acids 189-220) is involved in both functions [85, 107, 164, 214]. In the available structures, the C37 residues 196-225 are disordered [82].

The C53 subunit was further reported to interact with TFIIIC, highlighting its importance in transcription initiation (see section 1.3.4).

The Pol III enzyme is completed by the C31-C34-C82 heterotrimer, a subcomplex specific to Pol III. The heterotrimer sits on top of the clamp domain, and contains seven WH-domains. It is considered to be homologous to the Pol II general transcription factor TFIIE and the extended winged-helix (eWH)-domain of Pol I subunit A49, which bind their cognate RNAPs at equivalent positions. The presented similarities between the three RNAPs underline the notion that Pol I and Pol III have stably incorporated transcription-factor like proteins into the core enzyme [24, 201]. In agreement with this, the heterotrimer, and the two N-terminal WH domains of C34 in particular, are critically important for interaction with TFIIIB and for promoter melting [21] (see section 1.3.4 for more on this).

Since structures of all RNAPs and the Pol-II pre-initiation complex (PIC) have become available recently, the similarity can be appreciated in a side-by-side view (see Figure 1.2).

The cryo-EM structures of transcribing and unbound Pol III revealed that Pol III can adopt at least two conformations, which are characterized by different widths of the cleft and are referred to as closed-clamp and open-clamp states. In the DNA-free state, both conformations are sampled, whereas transcribing Pol III is locked in a closed-clamp conformation [82]. The structural changes are relayed via the stalk, which contacts the clamp domain and the heterotrimer. When the clamp closes, the stalk moves in the opposite direction as the heterotrimer.

In the cryo-EM map of transcribing Pol III, density for the DNA-RNA hybrid was weak

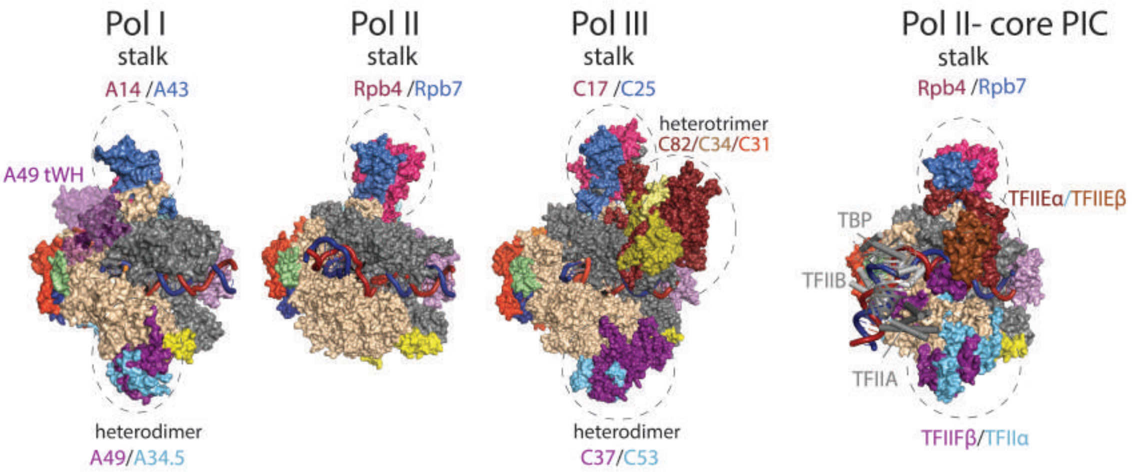


Figure 1.2: Comparison of Pol I, II and III, and the Pol II PIC reveals similarities between Pol III specific subunits and Pol II transcription factors. Homologous subunits are coloured identically, and the Pol I A49 tWH domain is rendered transparent to indicate its flexibility. Proteins in the Pol II PIC without counterparts in the Pol I and Pol III enzymes are shows as grey cartoons. Based on PDBs 4c3j, 5c4x 5fj8 and 5fyw. Adapted from [112]

compared to Pol I and Pol II maps, which might be important to allow factor-independent transcription termination, a feature unique to Pol III [82].

Transcription initiation by all known RNAPs requires auxiliary factors, known as TFs. TFs bind and recognize promoter elements and recruit their cognate RNAP, thereby determining the transcription start site (TSS). TFs not only position RNAPs, they are also essential for melting the DNA duplex and loading the template strand into the active site. The three eukaryotic RNAPs have evolved specialized TFs that can only recruit their corresponding RNAP, explaining why a given gene is transcribed exclusively by one RNAP and not another.

The Pol III system mostly relies on two well-characterized and conserved TFs, namely TFIIIB and TFIIIC, but certain genes like the 5s rRNA and the metazoan type III promoters require additionally TFIIIA and SNAPc, respectively. Interestingly, the TFs that are located close to the TSS and are in direct contact with Pol III appear to be much stronger conserved than those more distantly positioned [81].

1.3.2 *TFIIIB*

TFIIIB is the only transcription factor required at all Pol III genes and is sufficient for Pol III recruitment and transcription *in vitro* [99]. TFIIIB consists of three proteins that only stably associate when bound to DNA. The first component, TATA-binding protein (TBP), has a unique role in eukaryotic transcription as it is the only factor shared by all three RNAPs (although not essential in the Pol I system [110]). The second component, B-related factor (Brf1), is named due to homology of the N-terminal part of Brf1 with the Pol II transcription factor TFIIB. Lastly, subunit B" (B double prime (Bdp1)) is unique to the Pol III system.

Brf1 is a modular protein of 67 kDa with an N-terminal Zn-ribbon domain, followed by

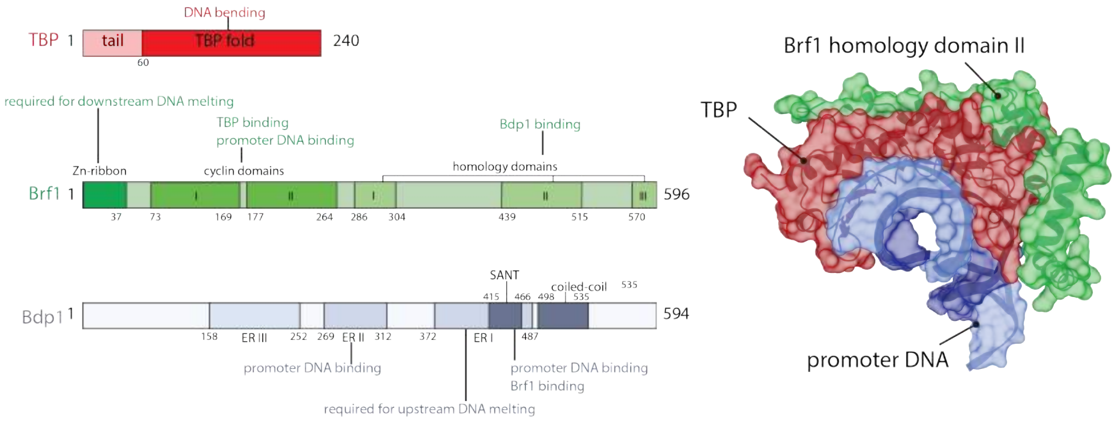

DNA bending
TBP 1
tail
TBP fold
240
60
required for downstream DNA melting
TBP binding
promoter DNA binding
Bdp1 binding
Zn-ribbon
cyclin domains
homology domains
Brf1 1
I
II
I
II
III
596
37
73
169
177
264
286
304
439
515
570
SANT
coiled-coil
535
415
466
498
535
Bdp1 1
594
158
ER III
252
269
ER II
312
372
ER I
487
promoter DNA binding
promoter DNA binding
Brf1 binding
required for upstream DNA melting
Brf1 homology domain II
TBP
promoter DNA

Figure 1.3: TFIIIB subunits and domains, annotated with information presented in section 1.3.2. The crystal structure of a fragment containing Brf1 homology domain II , TBP and promoter DNA is shown on the right (PDB 1ngm).

two cyclin folds—also found in TFIIB and Rrn3—and a large C-terminal extension [98]. Structures of the Pol II-TFIIB complex show that the Zn-ribbon protrudes in the Pol II active site cleft and that the cyclin folds interact with upstream DNA and the Pol II wall [118, 176]. In agreement with this model, Brf1 lacking the Zn-ribbon (Brf1 Δ1-68) is transcriptionally inactive, but transcription can be restored when an artificial DNA-mismatch close to the TSS is introduced [100].

The C-terminal part of Brf1 contains three "TFIIIB homology domains", identified by multiple sequence alignment [113]. The homology domain II (residues 439-515) is the central interaction site that glues together TFIIIB; in particular, a crystal structure revealing Brf1 fragment 437-506 bound to a TBP:U6 promoter complex showed how this short segment wraps around TBP [95]. The architecture revealed in this crystal structure inspired the design of a fusion protein in which the TBP core is placed in between Brf1 homology domains I and II, and this chimeric construct is fully active *in vivo* and *in vitro* [108]. A dense site directed mutagenesis study of the Brf1 homology domain II showed that this is also the principal anchorage site for Bdp1 [108], which has been confirmed by an NMR study pinpointing Brf1 residues 470-495 as the main interaction site with the SANT domain of Bdp1 [175].

At mammalian type III promoters, Brf1 is replaced by the isoform Brf2 [195] (see below and section 1.4.3 for more details).

The last component, Bdp1, has remained structurally more elusive until recently. While sequence-based conserved domain analysis using SMART [127] predicts a SANT (Swi3, Ada2, N-Cor, and TFIIIB) domain from residues 416 to 464 and a C-terminal coiled-coil domain, the remainder of the protein appears to be of low complexity. However, three essential regions (ERs) have been defined on a functional level; deletion of these regions are lethal in yeast [87]. The C-terminal ERI (residues 372–487) contains the SANT domain and flanking regions. Biochemical analysis have established that deletions within ERI (specificallyΔ355-372) render Bdp1 unable to support transcription initiation, but that this can be rescued by introducing mismatches at the upstream edge of the transcription bubble [100]. Furthermore, Bdp1 ERI crosslinks to the Pol III subunits C128 and C37 [84]. ERII (residues 269–312) was shown to crosslink to promoter DNA [186] and is required for TFIIIC-dependent transcription in vitro [120]. Footprinting and hydroxyl-radical probing experiments suggest that ERI and ERII are buried in the PIC, whereas the N-terminal ERIII (residues 158–252) is more accessible [87, 186].

The TFIIIB-DNA complex is unusually stable once assembled, and known to resist high concentrations of salt and heparin. This is a result of kinetic trapping of DNA within the TFIIIB complex and is only observed when all three components are present. The half-life of the TFIIIB-DNA complex is ~95 min *in vitro* [34]. This number is similar to the doubling time of exponentially growing yeast cells, suggesting that TFIIIB remains permanently associated with its promoter [34].

During the course of my PhD, crystal structures of truncated human Brf2-TBP and Brf2-Bdp1-TBP complexes bound to a short U6 promoter fragment were published [65, 66], shining light on the structural organization of TFIIIB. These confirmed the homology between TFIIB and Brf1 and revealed a similiar arrangement as seen in the Pol II PIC [76], where the cyclin folds and the Bdp1 SANT domain are bound on opposite sites of the DNA, bridged by TBP and the Brf1 homology domain II. However, these structures lack the functionally important residues in Bdp1 and Brf1 that are required for promoter opening.

Interestingly, Brf2 contains an oxidizable cysteine residue (C361) in a helical element termed 'molecular pin'. Oxidation of C361 abolishes Brf2-dependent transcription, making Brf2 a redox-sensing TF that couples expression of selenocysteine tRNAs to oxidative stress [65].

1.3.3 *TFIIIC*

TFIIIC is the largest Pol III transcription factor, with six subunits and a mass of 520 kDa. Insight from limited proteolysis experiments [137] and early rotary shadowing EM images [183] have established that it is organized in two subcomplexes, namely τA, comprising τ131, τ95 and τ55 and τB, comprising τ138, τ91 and τ60. τA binds the A-box promoter element and interacts with TFIIIB and Pol III upstream of the TSS, whereas τB binds the B-box promoter element with high affinity. Depending on the distance between the A-box and B-box, TFIIIC can adopt a compact globular shape when these elements are closely spaced, or a 'dumb-bell' like shape when A- and B-box are separated by more than 50 bp, where the two domains appear to be tethered by a flexible linker [183]. The structure of TFIIIC has been investigated by X-ray crystallography and crosslinking mass spectrometry in the lab over many years, and (partial) structures of all subunits are available, although coverage for τ138 is very poor.

1.3.3.1 *τA*

τ131 is the only component of TFIIIC that can be crosslinked to DNA upstream of the TSS [7]. τ131 contains a protease-sensitive N-terminal linker, an N-terminal tetratricopeptide repeat (TPR) array and a second putative TPR array located at the C-terminus. The crystal structure of the N-terminal TPR array has been solved and revealed the presence of ten TPR repeats, split into two clusters by a central region that causes a bend after TPR 6 [134]. This N-terminal part of τ131 is the central protein-protein interaction hub in TFIIIC and contains high affinity binding sites for Brf1 (K_D =300 nM [141]), Bdp1 and τB.

τA is completed by the τ95-τ55 dimer. DNA-photocrosslinking studies have positioned τ95 and τ55 on opposite sites of the A-box DNA helix [7], but only the τ95 orthologue was shown to possess (nonspecific) DNA binding activity in isolation [49]. τ95 contains an N-terminal dimerization domain which interacts with τ55, a linker that connects to a DNA binding domain, and a C-terminal acidic tail.

τ55 is a chimeric protein that contains a N-terminal histidine phosphatase domain (only found in hemiascomycetes) and a conserved C-terminal dimerization domain that interacts with τ95. Interestingly, the histidine phosphatase domain is not required for yeast viability or DNA binding or *in vitro* transcription, but deletion confers a growth defect under elevated temperatures or sub-optimal carbon sources. It has been suggested that it might couple transcription to metabolism [135]. In addition, the τ95-τ55 dimer exists independently of holo-TFIIIC in yeast [135]. Since the N-terminal histidine phosphatase domain is an active, *bona fide* enzyme, it is also possible that it might modify directly the transcriptional machinery in the nucleus [193].

A number of crystal structures are available for τ95 and τ55; the τ55 histidine phosphatase domain of *S. cerevisiae*, a structure of the τ95-τ55 dimerization domains, resembling the TFIIF-triple barrel fold, and a structure of the τ95 DNA-binding domain [134, 149, 192, 193]. The latter two were solved for the *S. pompe* proteins. Interestingly, the τ95 DNA-binding domain shows only micromolar affinity to both double-stranded and single-stranded DNA and little sequence specificity, as it can bind both A-box and B-box oligonucleotides *in vitro*, and contains a conserved acidic tail that auto-inhibits DNA binding [192], which was proposed to increase specificity of DNA binding.

1.3.3.2 *τB*

τB constitutes the module responsible for high affinity interactions with promoter DNA. The τ138 subunit is the largest subunit, and predicted to contain several DNA-binding domains; a high-mobility group (HMG) domain and the very N-terminus, and five WH domains dispersed throughout the protein. However, a large portion of τ138 is predicted to be disordered. Therefore, structural information about τ138 is limited to a crystal structure of a central eWH domain encompassing residues 546-641 [134]. Using site-specific incorporation of photo-reactive nucleotides, τ138 was mapped over the B-box [16], but the nature of these experiments does not allow to pinpoint the domain(s) in τ138 responsible for this interaction. However, a mutation in eWH-domain 3 (glycine 349 to glutamate) was reported to decrease affinity to DNA [126].

Subunits τ91 and τ60 form a stable dimer within τB, and τ91 has been mapped over the transcriptional end site of the 5s rRNA gene [16], although recombinant τ91 only has detectable affinity to single-stranded, but not double-stranded DNA [6]. The DNA-binding domain of τ91 is located at the N-terminus (residues 1-159) [149]. A crystal structure of τ60 in complex with τ91 (lacking the DNA-binding domain) has been solved, revealing that both proteins form similar seven-bladed β-propeller domains that stack perpendicular against each other. τ60 in addition contains a C-terminal α-β-domain [149]. The β-propeller domains of τ60 and τ91 are required and sufficient to form a scaffold for τ138 assembly, yielding a B-box binding τB module [49, 149].

Interestingly, recombinant τB is able to moderately stimulate transcription of TATA-box containing genes, but not TATA-less genes, perhaps through recruitment of TBP by the C-terminal domain of τ60 [49, 149].

While the preceding discussion was based on the *S. cerevisiae* proteins, it should be mentioned that the human τ60 homologue, TFIIIC90, contains histone acetyltransferase-

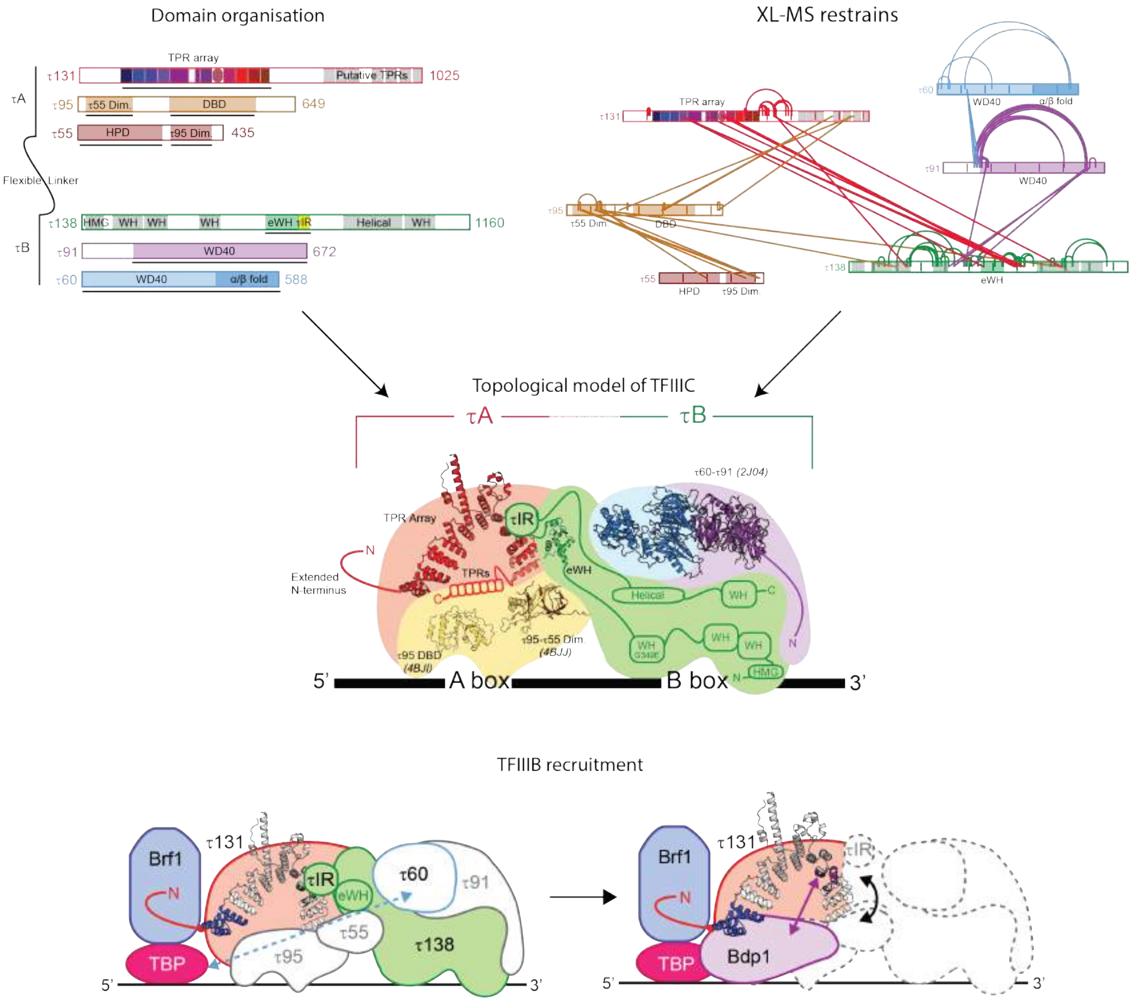

Domain organisation
TPR array
τ131
Putative TPRs
1025
τA
τ95
τ55 Dim.
DBD
649
τ55
HPD
τ95 Dim.
435
Flexible Linker
τ138
HMG
WH
WH
WH
eWH
τIR
Helical
WH
1160
τB
τ91
WD40
672
τ60
WD40
α/β fold
588
XL-MS restrains
τ60
WD40
α/β fold
TPR array
τ131
τ91
WD40
τ95
τ55 Dim.
DBD
τ138
eWH
τ55
HPD
τ95 Dim.
Topological model of TFIIIC
τA
τB
τ60-τ91 (2J04)
TPR Array
τIR
N
Extended N-terminus
TPRs
C
eWH
Helical
WH
C
τ95 DBD (4BJI)
τ95-τ55 Dim (4BJJ)
WH G349E
WH
WH
N
N
HMG
5'
A box
B box
3'
TFIIIB recruitment
Brf1
τ131
τIR
eWH
τ60
τ91
N
τ55
τ138
τ95
TBP
5'
3'
Brf1
τ131
τIR
N
TBP
Bdp1
5'
3'

Figure 1.4: Architecture of TFIIIC and interaction with TFIIIB. Top: TFIIIC subunits and domains, and distance restrains from XL-MS. Black lines under domains indicate available crystal structures from *S. cerevisiae* or *S. pombe*. Middle: Model of TFIIIC, based on available crystal structures and XL-MS restraints. Bottom: Recruitment of TFIIIB involves internal rearrangement of TFIIIC and replacement of the τIR from τ131. Adapted from [134], Figures 1 and 6.

activity, which might be important for re-shaping chromatin at Pol III genes and other TFIIIC loci (see section 1.5.2).

While the subunit composition of τA and τB is well established, the overall architecture of holo TFIIIC is more difficult to address. A genetic screen for suppressors of the previously mentioned τ138 G349E mutant, which impairs B-box binding, has revealed mutations in τ91 and τ131, suggesting direct interaction of τ138 with these subunits. However, due to the complex molecular interactions and the indirect read-out, it is equally likely that the suppresor mutants increase affinity of the τA module to the A-box or to TFIIIB.

By expressing different combinations of TFIIIC subunits in insect cells, Jourdain et al. demonstrated that the τA subunit τ95 can pull down the τB subunits τ138 and τ91, arguing for a role in bridging the two τmodules [94]. More recently, crosslinking-mass spectrometry of purified TFIIIC has yielded a more comprehensive map of subunit interactions [134].

This confirms the τ95-τ138 interaction, but crosslinks between τ95 and τ91 were absent. However, a large number of crosslinks were found between the N-terminal TPR array of τ131 and a central region of τ138, encompassing the eWH domain and an adjacent stretch of ~50 amino acids that are predicted to be disordered. This τ interacting region (τIR) is necessary and sufficient to bind the τ131 TPR array with high affinity (~100 nM). The binding site of the τIR was mapped on the TPR8 of τ131 [134].

Interestingly, the binding spot of the τIR in the τ131 TPR overlaps with the binding site of Bdp1, and binding of Bdp1 and the τIR are mutually exclusive.

Figure 1.4 summarizes the current structural knowledge about the architecture of TFIIIC.

1.3.3.3 *Assembly of the TFIIIB-TFIIIC complex*

The mechanism of TFIIIB recruitment by TFIIIC has been the focus of a number of studies, and proved to be a complex and dynamic process. Footprinting experiments using DNase I digestion of radioactively labelled tRNA probes have illustrated the sequential nature of the assembly pathway [105]. This revealed that the A- and B-box elements stay protected by TFIIIC throughout the assembly, but that the large TFIIIB footprint is only observed when both Brf1 and TBP are present. However, this complex remains labile. Addition of Bdp1 generates the very stable, heparin resistant TFIIIB complex.

Interestingly, addition of Pol III to assembled TFIIIB-TFIIIC renders the A-box susceptible to DNase I digestion, indicating dissociation of τA from its promoter element [102].

The first and limiting step of this assembly is the binding of Brf1 to the N-terminal part of τ131 [185]. This interaction involves both the protease-sensitive N-terminus and the TPR array of τ131, but the N-terminus is required for high-affinity interactions [128, 129, 134]. It was suggested that TFIIIC exists in an auto-inhibited conformation prior to Brf1 binding, because the binding affinity of smaller fragments of τ131 (N-terminus-TPR5, and TPRs 6-9) are higher then that of the entire TPR [143]. Moreover, dominant mutations in TPR2 increase the association of Brf1 with τ131 and result in increased transcription [144]. However, the mechanism of this auto-inhibition is unclear.

Yeast two-hybrid studies and competition experiments using short Brf1-derived peptides allowed to pinpoint regions in Brf1 that bind the τ131 TPR array, demonstrating that the Brf1 cyclin folds interacts with τ131 (Nterm-TPR5), and equally a fragment encompassing Brf1 homology domain II (residues 435-561) bind to the same τ131 fragment [128]. Interestingly, a peptide corresponding to the first α-helix of cyclin fold I in Brf1 releases the auto-inhibition of TFIIIC by binding to τ131 TPRs 6-9. Circular dichroism-spectroscopy experiments suggest that the binding of Brf1 to τ131 involves the formation of new α-helices [141].

Following binding of Brf1, TBP is recruited to the complex (although given the high affinity of Brf1 to TBP, a pre-assembled Brf1-TBP complex might also exist prior to TFIIIC binding) through binding sites in Brf1 (described in section 1.3.2) and contribution of the C-terminus of τ60 [49, 149].

Lastly, Bdp1 enters the complex, which involves major conformational changes. Bdp1 binding is driven by interactions with Brf1 and TBP (see section 1.3.2) and a high-affinity site in the τ131 TPR array, including TPR2 and TPR8 [129, 134]. As already mentioned, Bdp1 binding directly competes with a high-affinity intra-molecular TFIIIC interaction between τ131 and the τ138 τIR. Therefore, Bdp1 must displace the τIR region and might trigger a large-scale conformational change that renders the entire TFIIIB-TFIIIC complex competent for recruiting Pol III. This mechanism would serve as a proof-reading function to prevent recruitment of Pol III to incomplete TF complexes.

1.3.4 *Pre-initiation complex formation and promoter opening*

While TFIIIC is generally required for initial rounds of transcription, it is dispensable once TFIIIB is assembled at the promoter as demonstrated by heparin treatment of TFIIIC-TFIIIB-DNA complexes, which removes TFIIIC, but not TFIIIB [103]. In agreement with this, interactions between TFIIIC and Pol III are relatively sparse. The only reported interactions are between τ131 and the Pol III subunits C53 and ABC10α, based on yeast two-hybrid screens [50, 59].

The interactions between TFIIIB and Pol III are more numerous. Early on, binding of Brf1 and C34 was reported using two-hybrid screens [59, 206] and later confirmed by mutagenesis [21] and crosslinking studies [115]. Similarly, the N-terminal Zn-ribbon domain of Brf1 was shown to interact with the stalk subunit C17 [56, 115]. In fact, the architecture of the Pol III-TFIIIB complex was extensively probed using site-specific incorporation of non-natural photoreactive amino acids [85, 114, 115, 205, 213], with some of these studies published only after the beginning of this thesis. In summary, this suggested that Brf1 binds to Pol III similar as TFIIB does to Pol II, with additional interactions to the Pol III-specific heterotrimer subunits C34 and C82 [114, 115, 205].

Bdp1 was shown to interact with the C128 subunit via a functionally important, conserved region that is located directly N-terminal of its SANT domain [85]. Bdp1 also crosslinks to a hotspot in the Pol III heterodimer subunit C37, which also shows extensive crosslinks to C34, C128 and τ131 of TFIIIC [213] (see also Figure 1.1).

In addition to Pol III recruitment, TFIIIB also facilitates promoter opening; the two activities are separable because certain mutants can recruit Pol III but are transcriptionally inactive [71, 106]. The melting of the DNA duplex into single-strands could be followed by $KMnO_4$ oxidation of single-stranded thymines, showing that the upstream region of the transcription bubble melts at a lower temperature then the region around the TSS, presumably reflecting consecutive steps during promoter opening [102]. It was further shown that these two steps are coordinated by distinct TFIIIB domains, as a DNA template pre-opened at the upstream edge of the transcription bubble rescued the phenotype of the Bdp1Δ355-372 mutant, and a DNA template pre-opened at the downstream edge of the transcription bubble rescued the phenotype of the Brf1ΔZn-ribbon mutant. Photochemical crosslinking studies with wild-type and promoter opening-deficient TFIIIB mutants however did not reveal major differences in TFIIIB-DNA contacts, thereby arguing in favour of an allosteric role of TFIIIB in promoter opening, rather than through direct DNA contacts [104].

In the same study, the Pol III subunits C160, C128, C82 and C34 also crosslinked to the upstream edge of the transcription bubble, and these crosslinks were altered when TFIIIB mutants were used, suggesting that they are directly involved in promoter opening. This is in line with well-established roles of these subunits in transcription initiation [3, 21, 206].

1.3.5 *Transcription repression by Maf1*

Pol III transcripts are produced in large numbers and are therefore energetically costly to produce, consuming an estimated 15% of NTPs dedicated for transcription in growing yeast. Pol III is hence carefully regulated [145]. Pol III transcription is shut down under a variety of stress conditions, generally sensed through two signalling pathways in yeast - the Ras (Rat sarcoma) protein pathway, sensing glucose availability, and the TOR (target of rapamycine) pathway, which is a central regulator of metabolism and integrates nutrient availability and stress signals [32, 131]. Three effector proteins that directly bind the Pol III machinery and bring about transcription repression are well characterized; retinablastome (RB) protein, p53, and Maf1 [142, 208]. Of these, only Maf1 is found in yeast, whereas p53 and RB are only present in higher eukaryotes. While p53 and RB target all three RNAPs, Maf1 was discovered through its effect on Pol III and is required for tRNA repression by all tested stress signals (nutrient limitation, poor carbon source, oxidative and misfolding stress and more [142, 198]). However, human Maf1 was also reported to regulate Pol II transcription and inhibit expression of TBP, thereby indirectly downregulating global transcription by all three RNAPs [92]. Maf1 is conserved throughout eukaryotes and has extensively been studied in yeast, flies, worms, plants, parasites and mice [13, 211].

In yeast, Maf1 knockout increases total tRNA levels ~2.5 fold. This results in an increased translational fidelity by reducing translational termination read-through due to the higher abundance of the cognate terminator tRNA, but results in a temperature-sensitive growth defect on non-fermentable carbon sources [121].

The effects in higher eukaryotes are more profound, and demonstrate the central role of Maf1 and Pol III transcription for organismal physiology and metabolism. Whole body knockout of Maf1 in mice results in a striking phenotype of obesity resistance, caused by increased energy consumption through a futile tRNA synthesis and degradation cycle [14, 212].

Overexpression of Maf1 in mammalian cells on the other hand reduced anchorage-independent growth [92], a diagnostic for the tumorogenicity of cells. This result is consistent with the requirement of increased Pol III transcription for malignant transformation (see section 1.6.1).

Recently, it was demonstrated that overexpression of Maf1 in the gut of flies and worms increases the lifespan of these animals [57]. Intriguingly, this replicated the full effect of rapamycine treatment in adults, which is known to mimic a starvation state, and caloric restriction is known to improve health and extend lifespan in various animals, including primates [138]. It is highly likely that similar molecular mechanisms are active in monkeys and humans, underlying the potential health benefits of pharmacological intervention with the Pol III machinery.

It is noteworthy that Maf1 does not equally inhibit all tRNA genes, but rather that a set of housekeeping genes, which contains at least one tRNA for every amino acid, are less repressed [197]. The mechanism of this selective repression is currently unclear, but it has been suggested that highly active Pol III genes might be protected from Maf1, possibly through the mechanism of facilitated recycling.

1.3.5.1 *Mechanism of repression by Maf1*

During normal growth, Maf1 is phosphorylated and resides in the cytoplasm, however stress signals lead to rapid dephosphorylation and nuclear import [151] where it directly

binds to both Pol III and the Brf1 subunit of TFIIIB [41], and reduces occupancy of both factors at Pol III genes in yeast and humans [151, 163].
Interestingly, Maf1 mutants that lack phosphorylation sites do not constitutively repress Pol III, but only do so under stress conditions, indicating that phosphorylation of the Pol III machinery is equally important [142]. The many signalling pathways and post-translational modifications of the Pol III machinery in yeast are extensively reviewed in [31, 67]. In summary, these pathways regulate Pol III transcription by one or more of the following mechanisms: (i) inhibition of Pol III recruitment to chromatin, (ii) inhibition of the formation of a pre-initiation complex, most frequently by binding to the Brf1 and TBP components of TFIIIB (iii) modulation of the levels of active Pol III, TFIIIB or TFIIIC, either by post-translational modifications or modulation of total protein levels. All of these steps target the initiation step, which is probably due to the short size of Pol III genes which makes targeting elongation inefficient.
Maf1 contains three conserved sequence blocks, A, B and C, and an acidic tail. In *S. cerevisiae*, a non-conserved, unstructured linker is inserted between blocks A and B, which contains phosphorylation sites but can be removed to resemble the human Maf1 without affecting Maf1 function [142]. Crystal structures of human [202] and plant [188] Maf1 have been solved, and a low resolution (~20 Å) cryo-EM map of Maf1-bound Pol III was reported [202]. It was suggested that Maf1 achieves repression of Pol III through allosteric repositioning of the heterotrimer subcomplex which precludes binding of TFIIIB, explaining why binding of Brf1 and Maf1 to Pol III are mutually exclusive [202]. However, the published EM maps are of very low resolution, and interpretation was complicated by lack of an atomic structure of Pol III at the time. Moreover, to assign the Maf1 density, the authors compared maps of Pol III bound to a transcription scaffold to maps of Pol III-Maf1 without a transcription scaffold. Therefore, the observed rearrangement of the heterotrimer is likely to be caused by clamp closure of Pol III in the transcribing state, as observed in the high-resolution EM structures [82].

1.4 POL III PROMOTER TYPES

As already mentioned, distinct DNA sequences in the genome—known as promoters—demarcate Pol III genes and are recognized by the transcription machinery. Three distinct types of promoter sequences are known, and they require different sets of transcription factors. The three types are introduced in the following sections.

1.4.1 *Type I promoters - 5S rRNA*

The sole representative of the type I promoter type is the 5s rRNA gene *RDN5*. In yeast, 200 copies of the *RDN5* gene exist and they are organized in the ribosomal DNA (rDNA) locus together with the large, Pol I transcribed rRNA-precursor gene *RDN27* [181] in a tandemly repeated fashion. The rDNA locus constitutes the DNA component of the nucleolus, the nuclear site of ribosomal biogenesis. The type I promoter requires the TFs IIIA, IIIB and IIIC. TFIIIA is the archetypical Zinc-finger (ZF) protein and contains nine DNA binding Zinc-finger domains which recognize an extended internal control region (ICR) of the *RDN5* gene. Due to its high abundance in *Xenopus laevis* oocytes, TFIIIA has been extensively studied in the era preceding recombinant protein production (reviewed in [125]). While TFIIIA appears to be universally present in all eukaryotes, its primary sequence is remarkably less well conserved then the corresponding ICR. *S. cerevisiae* TFIIIA binds with its three N-terminal ZF domains to the ICR and recruits TFIIIC, an interaction that critically depends on the first ZF domain. Interestingly, TFIIIA has a dual function and binds large amounts of 5S rRNA in *Xenopus* Oocytes for storage which allows rapid ribosomal assembly during development.

Following TFIIIC recruitment, TFIIIB is assembled upstream of the TSS, forming the Pol III recruitment-competent TF complex. This process is described in detail in the following section, as it has been extensively studied at tRNA promoters.

1.4.2 *Type II promoters - tRNAs*

The type II promoter system is present at all tRNA genes, with the noteworthy exception of a single selenocysteine tRNA gene in metazoans (see section section 1.4.3). *S. cerevisiae* contains 275 tRNA genes, coding for 69 unique tRNA species, while the human genome contains 415 annotated tRNA genes and 176 tRNA-like genes [26].

Like the type I system, transcription of type II promoters depends on recruitment of TFIIIC downstream of the TSS, but binding is mediated through direct binding of conserved promoter elements by TFIIIC. These promoter elements are termed A-box and B-box and correspond to the D-loop and T-loop in the mature tRNA, respectively. The A-box is located 12-20 nt downstream of the TSS, and the B-box is located 30-60 bp downstream of the A-box; this large variability is caused by the fact that some tRNA genes (61 in yeast) contain introns. The TFIIIC module τB binds the B-box with high affinity and this interaction is the main determinant of selectivity and stability of TFIIIC binding to tRNA genes, whereas τA binding to the cognate A-box is of lower affinity. This is demonstrated by *in vitro* transcription of A-less or B-less promoters; the former support high levels of transcription but often at an altered TSS, whereas the latter supports low levels of transcripts initiated at the correct site, but requires high concentrations of template DNA and low ionic strength (reviewed in [62]). Therefore, the B-box determines promoter strength, whereas the A-box determines TSS-selection. In line with these findings, the consensus sequence of the A-box is more degenerate compared to the B-box sequence.

1.4.3 *Type III promoters*

Type III promoters are distinguished by the presence of strong upstream promoter elements and weak or absent downstream promoter elements. In yeast, the prototypic type III gene is the U6 gene, which contains a strong TATA box and can be transcribed without TFIIIC *in vitro* [90]. However, the yeast U6 gene still contains a weak A-box and a B-box uniquely positioned 130 bp downstream of the transcriptional end site and requires TFIIIC *in vivo* [23]. Other yeast tRNAs that contain a strong upstream U6-like promoter can also be faithfully transcribed *in vitro* without TFIIIC [42].

In contrast, vertebrates have evolved a truly distinct promoter type that is completely indepent of TFIIIC [147, 182] and contains no internal promoter element, but an additional "proximal sequence element" (PSE) located 50-70 basepairs upstream of the TSS, bound by an additional factor SNAPc, as well as a strong TATA box. Interestingly, the TATA box is the only element distinguishing vertebrate Pol II snRNA promoters from type III Pol III promoters, raising questions about how selective recruitment of Pol II or Pol III is achieved

(reviewed in [79]).
As already mentioned, Brf1 is replaced by Brf2 at the vertebrate U6 promoter and serves as a redox sensor.

1.5 POL III TRANSCRIPTION IN THE CONTEXT OF THE NUCLEUS

Pol III genes are unique genetic elements due to their small size and highly expressed nature. In the following sections, I will try to summarize to which extend these elements are occupied by the transcription machinery, and how they can influence expression of neighbouring genes and shape chromatin landscapes.

1.5.1 *Gene occupancy by the Pol III machinery in vivo*

chromatin immunoprecipitation (ChIP) studies of Pol III and Pol III TFs were performed in *S. cerevisiae* in the early 2000s [75, 146, 166, 189], leading to a comprehensive list of Pol III genes which had previously been assigned one at a time through the sensitivity of their expression to α-amanitin [109, 136, 210].
While the vast majority of Pol III genes identified in these studies were already known Pol III genes, the ability to quantify occupancy gave other interesting insights. For instance, all genome-wide studies consistently reported that essentially all Pol III genes were strongly occupied by Pol III in exponentially growing yeast cultures, confirming their high expression rates [75, 146, 166]. Moreover, Pol III genes were shown to be devoid of nucleosomes.
The studies also consistently reported highly correlated occupancies of TFIIIB and Pol III, but lower occupancy for TFIIIC. However, upon starvation Pol III occupancy was highly reduced, TFIIIB remained relatively stable, and TFIIIC occupancy drastically increased [166]. This suggests that TFIIIC and Pol III compete for the transcribed region. In recent work these finding were confirmed, and TFIIIC occupancy at tDNAs under stress conditions interpreted as an active mechanism of Pol III repression [33].
A complementary approach to study Pol III occupancy at its target genes used negative stain electron microscopy at the 5S rDNA gene (*RDN5*) locus [60]. It was possible to identify *RDN5* because of its association with large Pol I rDNA gene, which is easily identified in Miller chromatin spreads. Visual analysis of the 132 bp-long *RDN5* genes showed that they could be occupied by up to three polymerases, the highest theoretically possible Pol III density. Furthermore, given that the *RDN5* gene can be present in anywhere from 42 copies to 190 copies in different strains, the percentage of active genes, as well as the Pol III load per gene, were anti-correlated to the *RDN5* copy number. In all strains, the total number of *RDN5*-engaged Pol IIIs was similiar (72 ±4) showing that *RDN5* copy number was efficiently buffered. The study also estimated the transcription rate of Pol III at ~60 nt per second, indicating that transcription of the *RDN5* gene takes 2 seconds to complete and that a new polymerase initiates every 1.1 seconds.
Pol III occupancy is quite different in human cells. First, one study reported that the majority of Pol III-occupied sites in human cells are near SINEs rather than classical Pol III genes, albeit occupies near SINEs were lower then at classical Pol III genes [147, 210]. Second, only a fraction of all tRNA genes were occupied by Pol III, ranging from 30% [147] to 75% [153]. Comparison between four different cell types revealed that out of

the 513 predicted tRNA genes, a set of genes was consistently expressed among all cell types, likely to represent essential "housekeeping" functions, whereas a variable number of other genes showed cell-type specific expression, presumably matching the translational requirements of the respective cell type [153]. Both studies found that active Pol III genes are enriched in histone marks associated with open chromatin, and that active Pol III genes are frequently in the vicinity of Pol II genes and overlap with Pol II promoters[4] or enhancers [147, 153]. In addition, active Pol III genes showed enrichment for well known transcriptional activators characterized by their ability to activate Pol II transcription, even in cases where the Pol III genes were not close to Pol II promoters [153]. Strikingly, Pol III occupancy did not correlate with B-box strength. Therefore, chromatin accessibility seems to be the major determinant of expression of Pol III genes in higher eukaryotes, and Pol III genes and Pol II genes are extensively co-regulated. However, if Pol III transcription activates nearby Pol II genes or vice versa is unclear.

1.5.2 *Extra TFIIIC sites and tDNAs as insulators*

In addition to their crucial role in providing transcripts that power cellular growth, tRNA genes and the Pol III machinery play an important role in chromosome organisation in both yeast and mammals.

This was first discovered in *S. cerevisiae*, where deletion of a tRNA, which is located at the border of a transcriptionally silent chromatin domain, led to repression of a downstream gene [47]. Similar silencing of the downstream gene was reported when the A- and B-box of the tRNA gene or TFIIIC or TFIIIB were mutated. Therefore, engagement of the tRNA gene by the Pol III machinery is required to prevent spreading of heterochromatin, in agreement with the nucleosome-free nature of active Pol III genes [12]. The findings were later extended to human cells [161].

In the afore mentioned occupancy studies of Pol III factors in yeast and humans (see section 1.5.1), surprisingly a number of genomic sites that were bound by TFIIIC, but neither TFIIIB or Pol III were found [146, 166]. These sites were termed extra TFIIIC (ETC) sites. While only 8 ETC sites were found in *S. cerevisiae*, the ETC sites vastly outnumbered Pol III genes in humans (~5500 ETC sites vs ~1500 Pol III sites [147]).

It is now clear that both tRNA genes and ETC sites play important roles in establishing long-range chromosome interactions and nuclear domains (reviewed in [199]). In brief, especially the large genomes of animals are organized in domains that harbour clusters of active or inactive genes for co-regulation, and these domains are conserved through cell types and different species [58]. For instance, actively transcribed genes are clustered in discrete foci in the nucleus termed 'transcription factories' which can contain millimolar concentrations of Pol II, representing a 1000-fold enrichment [154]. Long-range interactions are facilitated by chromosome looping and underlie the enhancer-promoter interactions that regulate gene expression.

The most prominent insulator protein involved in looping is CTCF, only present in vetrebrates. TFIIIC shares many properties with CTCF; it can recruit condensin and cohesin complexes, which are believed to form rings around two DNA duplexes and thereby stabilize chromosome loops, and both TFIIIC and CTCF are enriched at boundaries of topologically associated domains ([199] and references therein).

---

[4] In cases where tDNAs overlapped with Pol II promoters, they were typically oriented in a divergent orientation so that Pol III would transcribe away from the Pol II gene

As one example of how TFIIIC can regulate Pol II gene expression, Crepaldi et al. demonstrated that association of TFIIIC to SINE elements regulates the formation of transcription factories upon synaptic activity of neurons, affecting dendritic growth and branching [37].

Apart from TFIIIC, TFIIIB can also influence transcription by Pol II. TFIIIB was shown to serve as a roadblock for both Pol II and DNA Pol δ, causing transcription termination or DNA replication termination [171, 172]. This was attributed to the unusual stability of TFIIIB on chromatin.

1.5.3 *Subnuclear localization of Pol III transcription*

Transcription by Pol III is highly coordinated in three dimensional space. For instance, the 5s rRNA genes are transcribed in the nucleolus [122]. While in yeast this localization is reflected in the linear arrangement of rRNA repeats in a single cluster on chromosome XII, the genes are more dispersed in higher eukaryotes and have to come together to form the nucleolus [69].

Interestingly, tRNA genes also dynamically associate with the nucleolar periphery, which requires active transcription and is only permitted for tRNA genes whose chromosomal location is distant from the centromer and telomers [9]. Transcription of tRNAs in the nucleolus is coupled to tRNA processing [83].

Coupling of tRNA transcription to tRNA export through tethering to nuclear pore complexes during the peak of tRNA synthesis in M phase has also been described [27].

1.6 POL III TRANSCRIPTION AND DISEASE

1.6.1 *Pol III and cancer*

Aberrant transcription by Pol III is associated with a number of surprisingly diverse disease phenotypes. The link between Pol I and Pol III transcription and cancer has attracted the most attention.

Already in 1896 enlarged nucleoli were observed in samples of human tumours [158], and nucleolar morphology is still used to diagnose and categorize tumour malignancy and prognosis; the larger the nucleolus, the worse the prognosis [40].

Enlarged nucleoli are a consequence of up-regulated Pol I and Pol III transcription to supply cells with enough ribosomes, as ribosome availability is thought to be the rate-limiting step in growth and proliferation [209]. Only recently it became clear that many of the oldest approved anti-cancer drugs, including DNA-intercalating platinum compounds, owe their therapeutic efficacy to inhibition of Pol I transcription, and new compounds are being clinically developed [22, 78].

Currently, there are no cancer drugs targeting Pol III, but several studies hint at a potential therapeutic window. Overexpression of components of the Pol III machinery is well documented in several tumor types, especially in liver and breast cancer. Both cancer types are linked to alcohol consumption, and it was shown that alcohol-induced overexpression of Brf1 results in increased Pol III transcription [215, 220, 221]. Tamoxifen prevents alcohol induced Brf1 overexpression in breast cancer cells and reduces cell proliferation [220]. While the association of upregulated Pol III transcription with cancers was long known, Johnson et al. showed that increased Pol III activity does not merely represent an adaptation of tumor cell metabolism to increased growth rates but is essential for the transformation process, and that downregulation of Brf1 levels in transformed cells decreases anchorage-indepent growth and tumor formation in mice [91].

A study on the vertebrate isoform Brf2 has recently provided unexpected insights into Brf2's mode of action and its link to cancer [65]. Brf2 overexpression has been associated with different lung cancer types, and is prognostic for poor survival rates. It also serves as a marker for diagnosis of these cancer types [130, 133]. Structural and functional work by Gouge et al. established Brf2 as a global redox sensing transcription factor, which uses a key cysteine residue to sense oxidative stress and couple it to selenocysteine metabolism and apoptosis [65]. The cysteine residue, C361, is located in a Brf2-specific short helical element, termed "molecular pin". This pin forms an interface with both TBP and DNA and is required for formation of a trimeric Brf2:TBP:DNA complex. Gouge et al further show that the local chemical environment makes C361 prone to oxidation, and that DNA binding is impaired when C361 is oxidized *in vitro*. Furthermore, induction of oxidative stress in cell lines specifically reduced levels of Brf2-dependent transcripts, including selenocysteine-tRNAs. Selenocysteine-containing proteins are primarily involved in reactive oxygen species (ROS) quenching, and a reduction in CeSys protein expression induces apoptosis. Therefore, oxidation of Brf2 indirectly triggers apoptosis by decreasing global levels of selenocysteine-tRNAs. Overexpression of Brf2 renders cancer cells insensitive to prolonged oxidative stress and allows them to proliferate rather than to undergo apoptosis, thus making Brf2 an important potential drug target.

1.6.2 *Pol III-associated developmental and neurological diseases*

While the involvement of Pol III transcription in cancer is not surprising giving its central role in proliferation, other diseases show cell-type or tissue specific phenotypes. Treacher Collins syndrome is a rare genetic disorder (prevelance of 1 in 50 000) which manifests itself by underdevelopment of facial bones and tissues [196]. It is caused by mutations in shared subunits of Pol I and Pol III, namely POLR1C, POLR1D, which encode for the human orthologs of yeast AC40 and AC19, which form a dimer and are both homologous to the bacterial RNAP α-subunit [38]. Most mutations are located on the interface of both subunits, but surprisingly do not affect dimerization when expressed in isolation. However, assembly into the holo enzymes of Pol I and Pol III is affected, resulting in fewer functional molecules and reduced transcription [204]. The disease phenotype probably manifests itself in development, where precursor cells of bone cells have very high protein synthesis requirements which cannot be met due to lack of ribosome and/or tRNAs [169, 204].

A second group of genetic Pol III-associated diseases known as leukodystrophies are characterized by severe neurological phenotypes, including developmental impairment and progressive de-myelation of neurons. They are associated with mutations in the two largest Pol III subunits, which leads to a decrease in Pol III levels in the brain [11]. Another disease characterized by decreased size of the cerebellum and intellectual disability was shown to be caused by mutations in Brf1, which leads to decreased occupancy of TFIIIB at tRNA promoters [15].

However, it still remains a paradox as to why somatic mutations in Pol III that decrease pre-tRNA transcription globally, cause cell-type specific phenotypes. One model put forward is that even if transcription levels of all tRNA genes are affected similarly, downstream processing such as splicing and base modifications lead to non-uniform changes in mature tRNA levels [4]. As a result, tRNA supply for certain codons is affected more strongly than for others, potentially causing translation infidelity and protein misfolding. Nerve tissue, which is known to be sensitive to folding stress, might hence be affected more strongly, leading to the devastating diseases.

1.6.3 *Pol III and links to immunity*

Several studies report links between Pol III and immunity. For instance, it was shown that NF-κB, a key transcription factor that orchestrates immune response, interacts with TFIIIB to increase tRNA levels which is required for high level cytokine production, which in turn stimulate growth of tumour cells at inflammation sites [68].
It was also shown that Pol III is the only RNAP that is present in the cytosol, where it acts as a DNA sensor for innate immunity. Pol III can transcribe foreign AT-rich DNA fragments that might be derived from viruses or bacteria, and the resulting RNA induced interferon expression which stimulates immune response [30].
Furthermore, mutations in Pol III subunits can interfere with this DNA sensing and render children and adults hyper-sensitive to viral infections [25, 152].

1.7 USING ELECTRON MICROSCOPY TO STUDY TRANSCRIPTION COMPLEXES

In recent years, cryo-EM has replaced X-ray crystallography as the method of choice for studying the molecular structures of large (>200 kDa) transcription complexes [36, 150]. This is driven by ongoing rapid advances in all aspects of cryo-EM, triggered by the introduction of direct electron detectors which offer unparalleled sensitivity and speed and have boosted the achievable resolution of biological samples from a domain-level description to a side-chain conformation-level of accuracy. Hardware improvements of the microscopes are being accompanied by sophisticated processing software, continuously improving processing speed, user-friendliness and achievable resolution.

Compared to X-ray crystallography, cryo-EM offers a number of advantages. First, it does not require crystals, which are notoriously difficult to grow for large, flexible biological samples. Second, sample requirements are typically orders or magnitude lower then for a crystallization project. Third, heterogeneous populations of particles can be classified *in silico*, allowing the determination of multiple structures from a single dataset. Forth, the cryo-EM experiment measures the phase of the scattered electromagnetic wave, and therefore does not suffer from the 'phase problem' encountered in crystallography. It is for that reason that low-resolution structures can be obtained by cryo-EM, whereas information obtained from X-ray crystallography is often useless unless crystals diffract to high resolution.

However, cryo-EM does not come without its own problems. In order to be visualized in an electron microscopes, the sample has to be embedded in an ultra-thin (~100 nm) film of vitrified ice at a concentration where individual particles do not overlap. Although this is often described as a 'native hydrated state', thin films do not represent the situation encountered in solution due to the extreme surface-to-volume ratios [64]. Particles confined in such a film have been estimated to collide 1000 times per second with the air-water interface [191]. Since this interface is hydrophobic, particles can either adsorb to the interface in a preferred orientation, dissociate into smaller components, or denature [64].

The typically low concentration of the sample applied to the EM grid further promotes complex dissociation when interactions are of low affinity.

The principle behind structure determination with electron microscopy lies in imaging particles from different angles and combining these different views into a three dimensional map that restores the three dimensional shape of the specimen. In single particle analysis projects, the different views are obtained by imaging many individual particles and assuming that each particle is identical to the next one. To obtain a faithful 3D reconstruction, particles must adopt different orientations in the sample relative to the electron beam. Due to adsorption affects, this can be difficult to achieve in practice. Another requirement is a sufficiently large signal-to-noise ratio in the data. Due to the strong damage inflicted on biological samples by the electron beam, the total electron dose must be carefully controlled and the signal-to-noise ratio is very low in cryo-EM. This further necessitates averaging of identical views, requiring tens of thousand to hundreds of thousands single particle images to obtain high resolution information, and is also the reason why only large samples (>300 kDa) have traditionally been amenable to the cryo-EM method. However, the size threshold is being pushed continuously (for instance [53]).

1.7.1 *The cryo-EM structure determination process*

The first step in the structure determination process is the preparation of a suitably thin sample containing the molecules of interest. For this, ~3 µL sample are pipetted on an EM grid mounted in forceps, excess liquid is removed by blotting with filter papers, and the grid is rapidly immersed into a coolant. These steps are usually performed in a dedicated instrument t (i.e. a VitroBot™) that has a humidity and temperature controlled chamber that allows maintaining the sample at the dew point to avoid evaporation during thin film formation, and has standardized blotting settings. Vitrification requires cooling rates of 10^6 K per second [64]. Suitable coolants are liquid ethane or propane which are cooled by liquid nitrogen to maintain them well below their boiling point.

EM grids are 3 mm across and manufactured from a metal support overlaid with a perforated carbon layer called "foil". Foils can contain irregularly shaped and sized holes ("lacey" foils), or holes of a fixed geometry (i.e. Quantifoil™ grids) which simplify automated data acquisition. Support layers can be added on top of the foils to provide a surface for the sample to adsorb to, which can be beneficial for concentrating the sample and/or depleting it from the air-water interface. The most common support layer is a thin (2 nm) film of amorphous carbon, which however degrades image contrast significantly and is only suited for imaging large samples.

Vitrified samples are transferred to an electron microscope and imaged. For high-resolution structure determination, data is collected in an automated fashion on a high-end microscope. Most crucial for image quality are the microscopes electron source—where field emission guns are state of the art and provide a coherent, bright, parallel and highly focused beam—the stage stability, and the detector. Direct electron detectors provide very fast read out times and high sensitivities. This allows to fractionate the total electron dose over several dozen movie frames while continuously exposing the sample to the beam. During data processing, movie frames are aligned to correct for motion during the exposure series. Sample motion is caused by mechanical stage drift as well as movement of the particles within the ice, which is caused by the energy transferred from the electron beam to the sample. Motion is hence corrected on a global level— aligning whole frames—and on a local level, either correcting motion within patches (as implemented in MotionCor2 [219]) or employing a particle-based tracking method [179]. Fractionating the total electron dose over several frames also allows for correction of radiation damage in a process called dose-weighting.

Motion correction is the first step of data pre-processing. In cryo-EM, images are collected with a defined defocus range (~-0.25 to -3 µm) to improve image contrast. This requires to correct for the microscopes contrast transfer function (CTF) in order to restore high resolution information. The CTF describes how the recorded images are distorted by the optical features of the microscope. The CTF is dominated by the optical abberation of the lense system, and by the defocus. The defocus can be estimated by calculating the fourier transform of the squared amplitudes (also known as the power spectrum) of the EM image, and fitting its maxima (called Thon rings), given that the spherical aberration of the lense, the acceleration voltage and the pixel size in the image are known. While the CTF was until recently estimated on a per micrograph level, software packages released in 2018 estimate local, per-particle defocus values (i.e Warp [194] and RELION 3 [222]), which is crucial to obtain reconstructions at better then 2 Å resolution.

Next, particles have to be identified in the CTF- and motion-corrected micrographs. A plethora of algorithms exist for automated particle picking, many of which are based on template matching. Care has to be taken not to introduce reference bias at this stage, as particles picked from random noise by matching to a reference structure will inevitably reproduce the reference structure. Template matching is outperformed by machine learning algorithms, as for example implemented in Warp [194].

In the last step of data-preprocessing, individual particle images are cropped out of the micrographs ('particle extraction').

Prior to attempting 3D reconstruction, the particle set is often cleaned from false-positives in a reference-free 2D classification step. This involves rotational and translational alignment of the particle projection images, and the resulting 2D averages show greatly improved contrast and are often indicative of data and sample quality and can help to validate 3D reconstructions due to their reference-free nature. However, particle cleaning based on 2D classification can lead to loss of images corresponding to rare views of the particles, as their corresponding 2D classes can appear noisy and poorly aligned.

To obtain a 3D map from the 2D projection images, the orientation of each imaged particle must be determined. In practice, this is achieved by comparing each image with projections of a reference structure. Reference structures are usually low-pass filtered to ~60 Å to avoid excessive reference bias. If the particle image set is sufficiently homogeneous, reference structures can now be determined directly from the data using stochastic gradient descent algorithms (first implemented in the CryoSparc package [160]). However, the handedness of such reference structures is unknown, and must be validated. When sufficient resolution can be obtained, this can be determined from the handedness of α-helices in the reconstruction of from fitting known structures into the map, if available.

To obtain high resolution 3D reconstructions, most datasets have to be extensively classified *in silico* in order to obtain sets of particles in the same conformation. Classification can further be focused on certain regions of the 3D map by providing a mask, which can be a powerful tool to separate particles that are bound to an interaction partner from those that are not, or to separate distinct conformations of a specific domain.

Once a suitable set of particles has been identified, final computational refinement steps can be employed to boost the resolution and/or map interpretability. In RELION 3, this involves per-particle defocus and beam tilt estimation, as well as particle-based motion correction and dose-weighting. Final maps are then postprocessed, a step in which the map is sharpened with a B-factor in order to counteract high frequency contrast fall-off. B-factors can be automatically determined from the data when the reconstructions yields maps at better than 10 Å[168]. Alternatively, when an atomic model is available and can be refined against the map, map-sharpening B-factors can be computed from the model coordinate B-factors in a local manner [88]. This is particularly helpful when the local resolution differs strongly across the map. Finally, the local resolution of the 3D map should be computed.

1.7.2 *Typical workflow of a single particle cryo-EM structure determination project*

1.7.2.1 *Complex reconstitution*

Before EM characterization can be started, the sample of interest has to be produced and purified. For the samples that are subject of this thesis, this requires purification of the constituent components of the complex of interest and establishing conditions under which they form a functional complex. Complex formation can be assayed in a number of ways, but size exclusion chromatography (SEC) using analytical columns that are compatible with low amounts of sample is often the method of choice. SEC of complex mixtures has the advantage of further purifying complex mixtures to enrich the complex of interest. However, SEC results can be misleading, as co-elution does not always mean binding. In addition, low-affinity complexes tend to dissociate during SEC.

1.7.2.2 *Negative stain EM*

Negative stain EM is well suited for initial characterization, as it immediately gives information about sample homogeneity and aggregation state, and allows the reconstruction of low resolution (up to ~15-20 Å) 3D maps that can give valuable biological insights and confirm that the complex of interest is present. Negative stain sample preparation involves applying the sample at low concentrations (~20 µg/mL) to carbon-coated copper grids, washing away excess liquid and staining with a heavy metal salt solution, typically uranyl acetate. Sample preparation is quick, and due to the high contrast, grids can be screened on low-end microscopes. However, in practice, image quality depends on the quality of the stain and it can be challenging to discern whether a sample is of low quality because of the biological sample or if it is poorly stained.

If the stain and particle concentration, distribution and homogeneity appear promising, datasets can be collected. In ideal cases these give rise to a low-resolution 3D map of the complex. 2D class averages can also be sufficient to judge sample quality or determine if a complex has formed or not.

1.7.2.3 *Screening cryo-EM conditions*

Typically, optimization of the freezing and sample preparation conditions is the bottleneck in the structure determination pipeline, and might require testing of various grid types (i.e with and without carbon support, different hole sizes and geometries), sample concentration, screening of buffer conditions, chemical crosslinking et cetera before a suitable sample is obtained. The ideal sample shows a high density of non-overlapping particles, random orientations of the particle in the ice, ice that is only marginally thicker then the particle diameter, and good image contrast. Only image contrast and particle distribution can be judged based on manual inspection of screening images. It is therefore best to collect small initial datasets to judge whether the complex of interest is present in the ice or if it has dissociated during sample preparation, and if sufficient angular coverage can be achieved. Therefore, obtaining a first 3D cryo-EM map is a milestone, even if it is of low resolution.

1.7.2.4 *Automated data collection and pre-processing*

Before setting up automated data acquisition, suitable areas of thin ice have to be identified by manual screening. It is noteworthy that different areas on the same grid can differ dramatically in terms of particle count and overall quality. When suitable regions have been identified, the user must decide on acquisition parameters, i.e. pixel size, defocus range, total dose, number of movie frames, camera mode, beam size and number of images per hole.

Ideally, data is (pre-)processed on the fly. Since 2018, the Warp package [194] is available which performs motion- and CTF-correction, particle picking and extraction with very little lag time. This gives valuable live feedback about the accuracy of the microscopes focusing procedure and the particle count, and can inform the user to skip the current acquisition area. Particles can also be periodically fed into processing software to compute 2D or 3D classes, which can help to determine if the sample meets specific criteria, i.e. presence of an interacting partner.

1.7.2.5 *Data processing*

While the basic steps of data processing are described in section 1.7.1, the choice of a classification strategy is critical for the final result of structure determination. Perhaps the most critical user decisions are the particle picking program, the choice of masks applied during classification, and whether or not to use 2D classification as a sorting step. Coaxing the highest resolution from the data and finding the most homogeneous sets of particles can be a lengthy process, and many strategies are often tested. Depending on the aim of the project—i.e obtaining sufficient resolution to enable *de novo* model building, or describing all conformational states in the data—it can be beneficial to either restrict classification to a minimum of different classes, which would boost resolution in the invariant parts of the structure but might deteriorate it in mobile parts, or to find the maximum number of distinct classes. Reconstruction algorithms such as the RELION MultiBody refinement program are also becoming available, and can help to maximize resolution in mobile areas while simultaneously describing the motion of individual bodies relative to each other by reconstructing partial densities while storing the relative orientation to other partial densities.

1.7.2.6 *Model building and validation*

The ultimate aim of cryo-EM structure determination is the interpretation of the reconstruction in terms of an atomic model. Depending on the resolution achieved, this can range from fitting of (homology) domains over construction of a backbone-model to *de novo* building of an all-sidechain model using programs like COOT [51]. Model building is still a very manual process which requires a lot of time and expertise. For an accurate backbone model, the resolution should exceed 4.5 Å, otherwise β-strands cannot be resolved [187]. Model building and refinement involves iterating between manual adjustment of the model (i.e. in COOT) and real-space refinement (i.e. in Phenix [2]). Real-space refinement is typically coupled to model geometry validation and density fit analysis programs, such as MOLPROBITY [28]. The real space refinement program aims to balance real-space density fit with satisfying physical restraints, i.e. bond length and angles, rotamer conformations, peptide dihedral-angel combinations (i.e. Ramachandran outliers), steric clashes between neighbouring atoms, and more. The relative weight of density fit and geometry optimization depends on the resolution of the map that is used as a refinement target. It is also helpful to monitor how secondary structure geometries are affected by the real-space refinement step, as distortion of helices or loss of hydrogen bonding in beta-sheets during refinement can be indicative of problems in the model, i.e. sequence register shifts. Finally, the fit of the model to the map can be evaluated globally by computing fourier shell correlation between simulated model maps and the experimental map.

1.8 AIMS AND MOTIVATION OF THIS THESIS

As pointed out in section 1.3.5, Pol III transcription is regulated at the level of initiation, and both stimulatory and repressive signaling pathways exist [31]. Misregulation of Pol III activity is of significant medical interest as over-activation of Pol III is a critical event in malignant transformation and Pol III transcription is associated with an increasing number of other phenotypes, from neurodevelopment to ageing. Yet, high-resolution structural information about the molecular interactions that regulate Pol III is lacking. In particular, three protein complexes govern Pol III activity: the general Pol III transcription initiation factors TFIIIB and TFIIIC, and the global repressor Maf1, which integrates signals from various cellular signaling pathways. Currently, structural information of TFIIIB is limited to homology models and crosslinking data, as well as a crystal structure revealing the architecture of TBP in complex with a 71 amino acid Brf1 fragment. No atomic resolution information about the interaction of TFIIIB with the Pol III specific complexes (i.e. the C82/C34/C31 heterotrimer) is available, and the structure of Bdp1 is largely unknown. While many individual crystal structures of domains and subunits of TFIIIC have been solved, it is still unclear how they arrange in the context of holo-TFIIIC, how they bind DNA and recognize the promoter DNA and how they recruit TFIIIB.
On the other hand, the crystal structure of a truncated human Maf1 construct has been solved. However, information about the Maf1:Pol III complex is limited to a 20 Å cryo-EM structure which only allowed the approximate positioning of Maf1 on Pol III, but its orientation remains unclear, impeding any mechanistic conclusions about the mode of action of Maf1. Hence, this thesis aims to answer the following questions:

I)

- What is the precise architecture of the TFIIIB:Pol III complex?
- Which Pol III specific elements are involved in the interaction?
- How does TFIIIB allow Pol III to initiate transcription? How do the different TFIIIB components interact with each other and with DNA to achieve melting of the DNA and transcription start site selection?

II)

- How are the TFIIIC subunits organized to form the plastic holo complex that accomodates deferentially spaced promoter elements?
- How is TFIIIB bound? What conformational changes occur during TFIIIB recruitment?
- How is transcription stimulated by TFIIIC?

III)

- What is the precise mode of binding of Maf1 to Pol III?
- Do Maf1 and TFIIIB use the same surfaces for binding to Pol III?
- How does Maf1 achieve repression of Pol III?

By answering these questions I hope to draw conclusions about the evolution of transcription initiation of eukaryotic RNAPs and reveal the adaptations that allow Pol III to undergo the rapid rounds of re-initiation that are unique to Pol III. Moreover, identifying interaction surfaces between Pol III and TFIIIB as well as Pol III and Maf1 at high resolution has the potential to inform structure-based drug design and might aid in the development of pharmaceuticals that can combat diseases caused by or linked to Pol III overactivity, such as cancer.

Part II

RESULTS

2

STRUCTURAL MECHANISM OF PROMOTER OPENING IN THE RNA POLYMERASE III PRE-INITIATION COMPLEX

2.1 RECONSTITUTION OF THE TFIIIB-POL III COMPLEX

2.1.1 *Expression and purification of TFIIIB subunits*

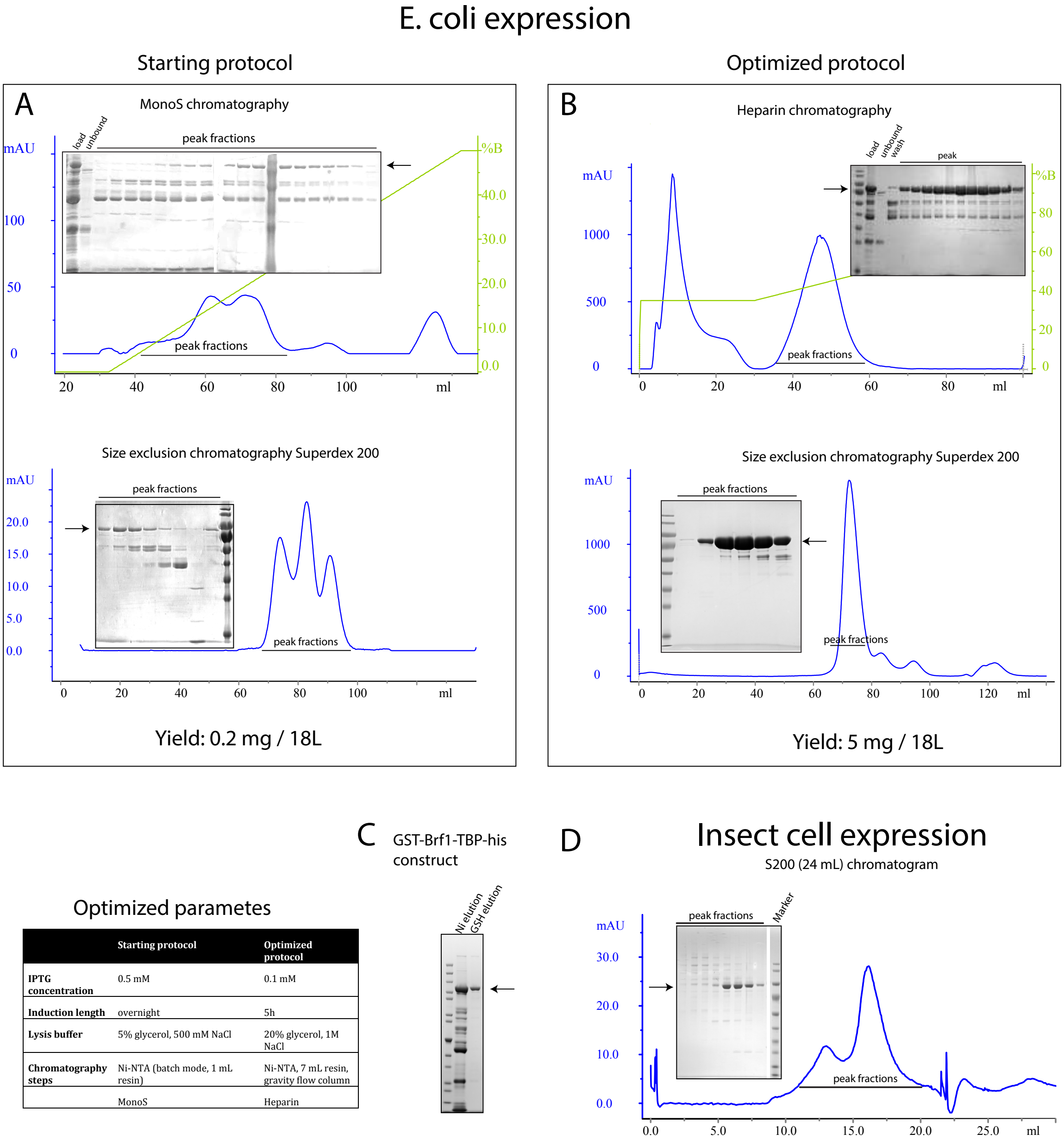
A prerequisite of all structural studies is the availability of sufficient amounts of the biomolecules to be investigated, which should be as pure and homogenous as possible. The Pol III purification is well established in the lab and based on large scale yeast fermentation and purification via IgG pulldown, tag cleavage and polishing via a MonoQ 10/100 GL column.

While previous efforts in the lab to purify and reconstitute TFIIIB showed that TBP and Bdp1 can be overexpressed and purified in *E. coli*, Brf1 purification was suffering from very low yields and strong degradation. Moreover, attempts to reconstitute TFIIIB from purified components or via co-expression were unsuccessful, as either no stable complex was obtained (when Brf1 and Bdp1 were incubated) or proteins precipitated (when Brf1 and TBP were incubated). Therefore, I focused my efforts on a fusion protein of Brf1 and TBP in which a TBP core is placed in between the N-terminal and the C-terminal portions of Brf1. This fusion construct, hereafter referred to as Brf1-TBP, has been shown to replace Brf1 function *in vivo* and *in vitro* [108]. Brf1-TBP had been cloned by a previous member of the lab with a C-terminal hexa-histidine (his)-tagged construct for expression in *E. coli*. Initial purifications gave very poor yields (0.2 mg from 18L of culture) and the protein was strongly contaminated with degradation products (see Figure 2.1). To overcome these limitations, three strategies were devised to obtain sufficient amounts of full-length Brf1-TBP

- Insect cell expression of Brf1-TBP using the bacculovirus system
- Cloning of a GST-Brf1-TBP-his construct, allowing the purification of full-length protein from *E. coli* by exploiting the affinity tags at both N- and C-terminus
- Optimization of expression and purification of the original construct

All approaches were pursued in parallel. While insect cell expressed Brf1-TBP did not suffer from degradation, yields were low and there were several impurities. The GST-Brf1-TBP-his construct was purified over Ni^{2+}-NTA and GSH beads and yielded fusion protein of high purity, however, precipitation of Brf1-TBP was observed upon removal of the GST tag. However, I could optimize purification of the original construct by implementing the alterations summarized in Figure 2.1 which increased the yield 25 fold while achieving much more homogeneous protein.

Bdp1 was also expressed in *E. coli*, initially as an N-terminal his-tagged protein, but later on as a N-terminal his-GST fusion protein with subsequent removal of the solubility tag, which gave ~10 fold higher yields.



	Starting protocol	Optimized protocol
IPTG concentration	0.5 mM	0.1 mM
Induction length	overnight	5h
Lysis buffer	5% glycerol, 500 mM NaCl	20% glycerol, 1M NaCl
Chromatography steps	Ni-NTA (batch mode, 1 mL resin)	Ni-NTA, 7 mL resin, gravity flow column
	MonoS	Heparin

Figure 2.1: Optimization of Brf1-TBP expression and purification. The band corresponding to full-length Brf1-TBP is indicated with an arrow.

2.1.2 *Reconstitution of the TFIIIB-Pol III complex*

To reconstitute the Pol III PIC, purified protein components were mixed with pre-annealed DNA oligos (see section 6), incubated and analysed by SEC, using a analytical Superose 6 column on an Äkta micro setup. The DNA scaffold was based on the U6 promoter, due to its independence of TFIIIC for *in vitro* transcription.

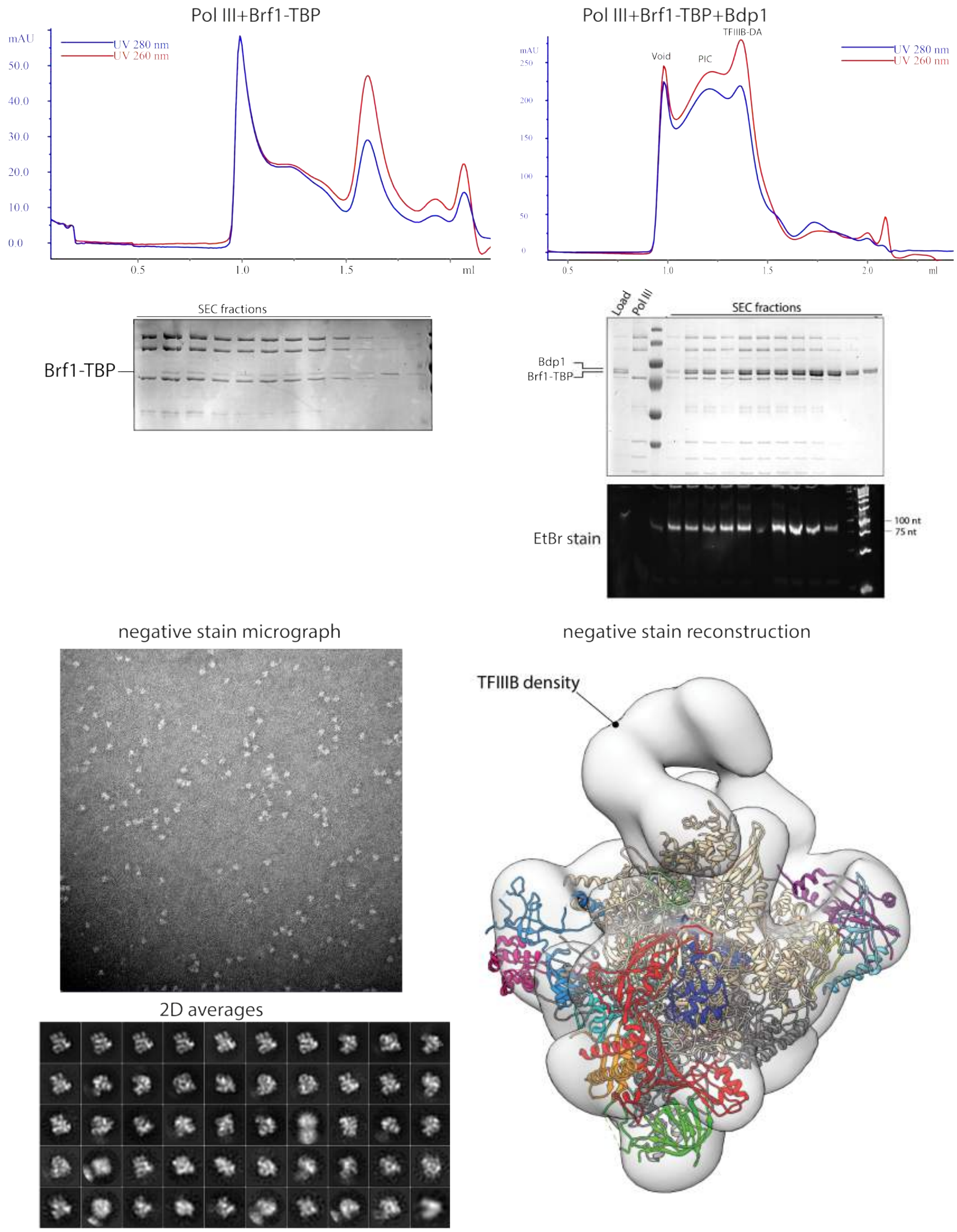


Figure 2.2: Reconstitution of the Pol III-TFIIIB complex. Top left: Substoichiometric binding is observed in absence of Bdp1. Top right: Under otherwise identical conditions, a stable, stoichiometric complex is formed when Bdp1 is included. Presence of all subunits and DNA is confirmed by coomassie and ethidium bromide staining, respectively. Bottom: Negative stain reconstruction of the open Pol III-PIC. A representative micrograph and reference-free 2D class averages are shown. Right: Pol III model fitted into negative stain density obtained for the TFIIIB-Pol complex.

Initially, attempts were made to reconstitute a complex between Pol III and Brf1-TBP to obtain a minimal PIC, but only substoichiometric binding was observed. However, when Bdp1 was included, a stoichiometric complex was obtained. SEC analysis showed a shift of the retention volume relative to Pol III:DNA and SDS-PAGE showed stoichiometric amounts of Brf1 and Bdp1 in the main peak (see Figure 2.2). The presence of DNA in the complex was indicated by a strong absorbance at 260 nm (A280/A260 =1.2) and by ethidium bromide staining of acrylamide-sucrose gels (Figure 2.2).

The sample was used to prepare negative stain grids and a 3D reconstruction revealed presence of additional density corresponding to TFIIIB (Figure 2.2). Indistinguishable negative stain reconstructions were obtained from samples with fully complementary DNA strands or for a scaffold mimicking the initially transcribing state.

The activity of this preparation was confirmed using *in vitro* transcription assays with different templates. Both plasmid DNA and the closed 80-mer DNA scaffold could be transcribed.

2.2 CRYO-EM OF THE POL III-TFIIIB COMPLEX

2.2.1 *Optimization of cryo-EM sample preparation of the TFIIIB-Pol III complex*

Initially, cryo-EM samples were prepared as for the Pol III structures [82], using SEC purified sample at a concentration of ~0.2 mg/mL on holey Molybdenum grids with a Vitrobot Mark II plunge-freezer. However, analysis of a dataset collected on a Titan Krios comprising 1462 micrographs revealed that the complex had dissociated during grid freezing, as evident by lack of TFIIIB density in the 3D reconstructions. Additionally, small dots were visible in the raw micrographs and in 2D classes, which I speculate represent free DNA duplexes oriented perpendicular to the ice surface.

In order to stabilize the PIC for cryo-EM, I decided to crosslink the sample with glutaraldehyde. During initial trials I identified a number of hurdles:

- The buffer used in our previous work containing $(NH4)_2SO_4$ was not suitable, as NH_4^+ showed reactivity with glutaraldehyde,
- Crosslinking led to the formation of large soluble aggregates, eluting in the void volume of the SEC column and greatly diminishing yields of monomeric sample.
- Crosslinked sample showed poor spread on EM grids, with particles accumulating in the center of the hole, and large variation of particle numbers across a grid.

Optimization of the protocol required to consider three parameters: yield of (momomeric) sample, crosslinking efficiency, and particle spread, and involved screening of crosslinker and sample concentrations, temperature and duration of the reaction, and changing of the vitrification device and EM grid type (see Figure 2.3).

During the process I acquired two preliminary datasets, which confirmed that glutaraldehyde crosslinking successfully stabilized the PIC (Table 2.1). The optimized protocol (see section 6) allowed me to obtain three large datasets[1], containing a total of 12522 micrographs, with the Pol III PIC bound to two different DNA scaffolds. The first scaffold mimicked the initially transcribing state with a short RNA in the active site, and contained

---

[1] cryo-EM data was collected with the help of Wim Hagen

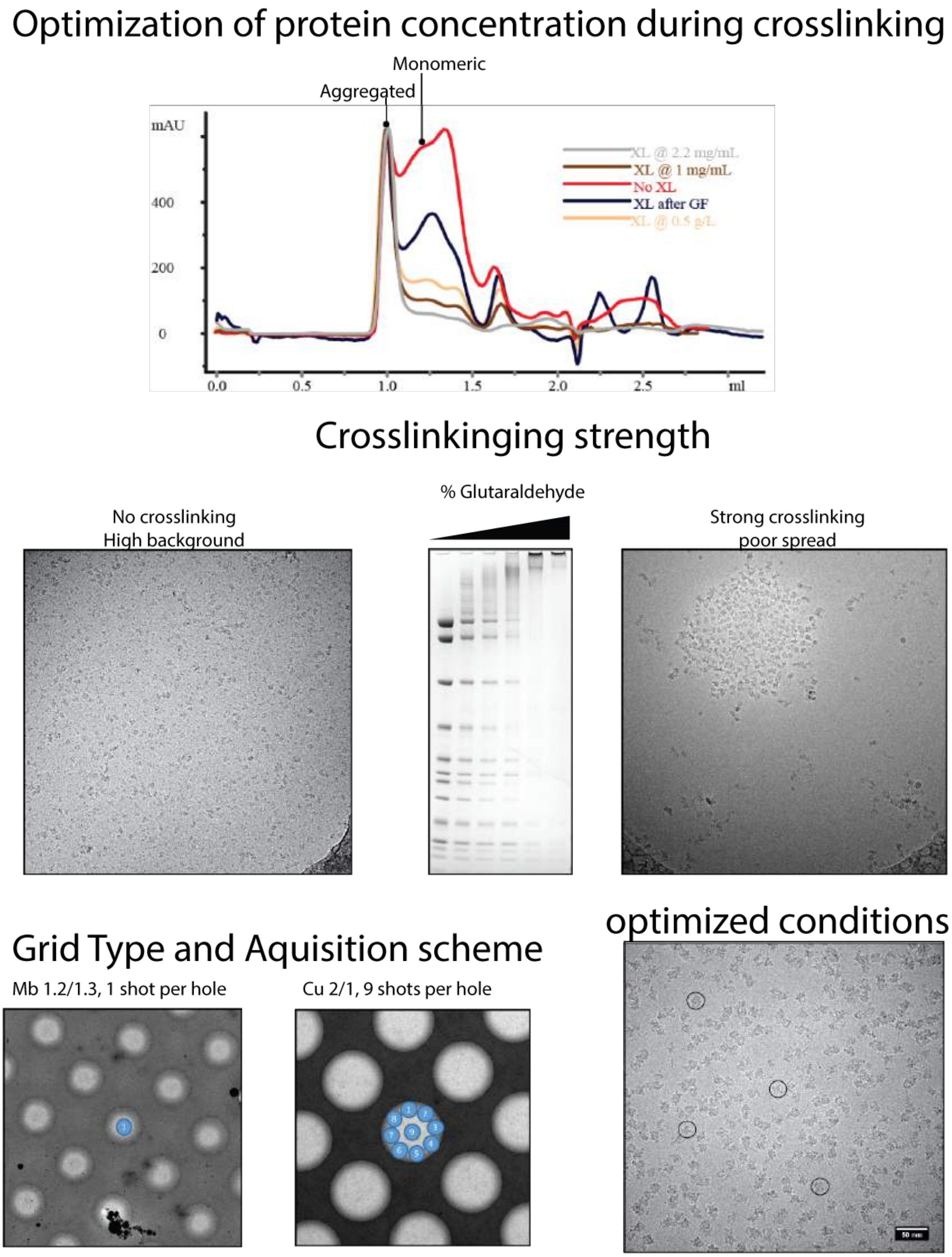

Optimization of protein concentration during crosslinking
Monomeric
Aggregated
mAU
400
200
0
0.0
0.5
1.0
1.5
2.0
2.5
ml
XL @ 2.2 mg/mL
XL @ 1 mg/mL
No XL
XL after GF
XL @ 0.5 g/L
Crosslinkinging strength
% Glutaraldehyde
No crosslinking
High background
Strong crosslinking
poor spread
Grid Type and Aquisition scheme
Mb 1.2/1.3, 1 shot per hole
Cu 2/1, 9 shots per hole
1
2
3
4
5
6
7
8
9
optimized conditions
50 nm

Figure 2.3: EM sample optimization of the Pol III PIC.Top: dilution of the protein sample prior to crosslinking reduced aggregation. Middle: While uncrosslinked sample dissociates during plunge freezing, highly crosslinked, SEC purified sample clusters in the center of the whole where the ice is thinnest. Bottom left: Changing of the grid type and acquisition scheme allows to obtain much more data. Bottom right: Micrograph under optimized conditions.

Table 2.1: Cryo-EM datasets acquired for the Pol III-TFIIIB complex

Sample	Closed complex	ITC	ITC	Closed complex (2 datasets)
Microscope	Polara	Titan Krios	Titan Krios	Titan Krios
Detector	Falcon I	K2	K2	K2
XL protocol	After SEC	Before SEC	Before SEC	Before SEC
# micrographs	699	1056	3872	8650
# of particles	19k	18k	29k	62k/18k/34k
Resolution (Å)	10.5	5.6	4.3	3.7/4.2/5.5
Dose (e-/Å^{-2})	48	90	61.3	60.0/61.8

a mismatch from base pair (bp) -7 to + 8 relative to the transcription start site and a 6-nt RNA oligonucleotide complementary to + 1 to + 6. The second scaffold was a fully complementary, closed DNA duplex. I obtained high resolution cryo-EM maps of the Pol III-PIC in four distinct states (see chapter 6 for details). These results have been published as an article in Nature [203]. Some of the following sections and sentence have been taken from this paper[2].

2.2.2 *Data processing and particle classification*

The cryo-EM data was extensively classified in order to separate distinct functional states (summarized in Figures 2.4 and 2.5, and described in detail in section 6.3.3). The classification strategy allowed me to obtain four distinct reconstruction of Pol III-TFIIIB, ranging from 3.7 to 5.5 Å resolution and comprise the PIC bound to a DNA:RNA scaffold (initially transcribing complex (ITC)), a spontaneously formed open DNA complex (OC), and two reconstructions of a closed DNA complex (CC). The maps are depicted in Figure 2.6.
During data classification, it was surprising to see that focused classification on downstream DNA in the ITC and CC datasets also gave rise to reconstructions which show OC-like upstream DNA but lacked downstream DNA (OCΔdownstream, Figures 2.4 and 2.5). This might suggest the existence of an initiation intermediate in which downstream DNA is mobile and only becomes ordered in a later stage of the transcription cycle, as it has also been described for bacterial RNA polymerase [54]. An alternative explanation for the lack of density in the OCΔdownstream map derived from the CC scaffold lies in the pseudo-symmetric nature of the U6 TATA box [95]. Since I used a wild-type promoter sequence, a subset of particles may have bound the DNA scaffold with reverse polarity, making the "downstream" DNA too short to be stabilized in the cleft. However, I favor the former explanation, as I also observe OCΔdownstream in the ITC dataset where the polarity is defined by incubating Pol III with the pre-formed transcription bubble.

2.2.3 *High resolution cryo-EM maps of the TFIIIB-Pol III structure in different functional states*

The ITC scaffold yielded a single conformation resolved at 4.3 Å, although a large fraction of particles showed poor or no density for downstream DNA and RNA and were discarded

---

[2] These parts were also originally written by myself

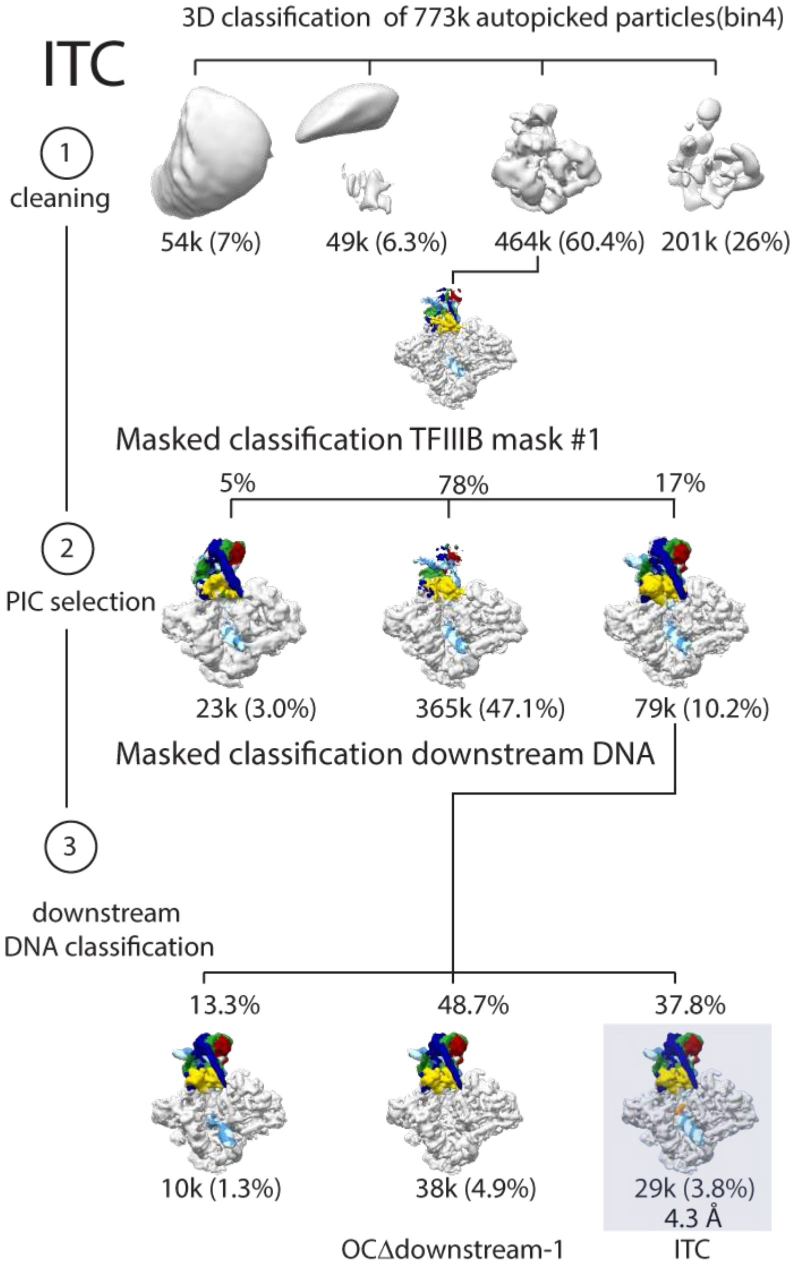

3D classification of 773k autopicked particles(bin4)
ITC
1
cleaning
54k (7%)
49k (6.3%)
464k (60.4%)
201k (26%)
Masked classification TFIIIB mask #1
5%
78%
17%
2
PIC selection
23k (3.0%)
365k (47.1%)
79k (10.2%)
Masked classification downstream DNA
3
downstream
DNA classification
13.3%
48.7%
37.8%
10k (1.3%)
38k (4.9%)
29k (3.8%)
4.3 Å
OCΔdownstream-1
ITC

Figure 2.4: Classification of particles from the ITC datset. The ITC dataset was cleaned (step 1) by classification in 3D (binned 4 times) and PIC particles were selected using a mask on TFIIIB (step 2). Finally, classification using a mask on downstream DNA yielded the ITC and the OCΔdownstream-1 maps (step 3). Figure taken from [203].

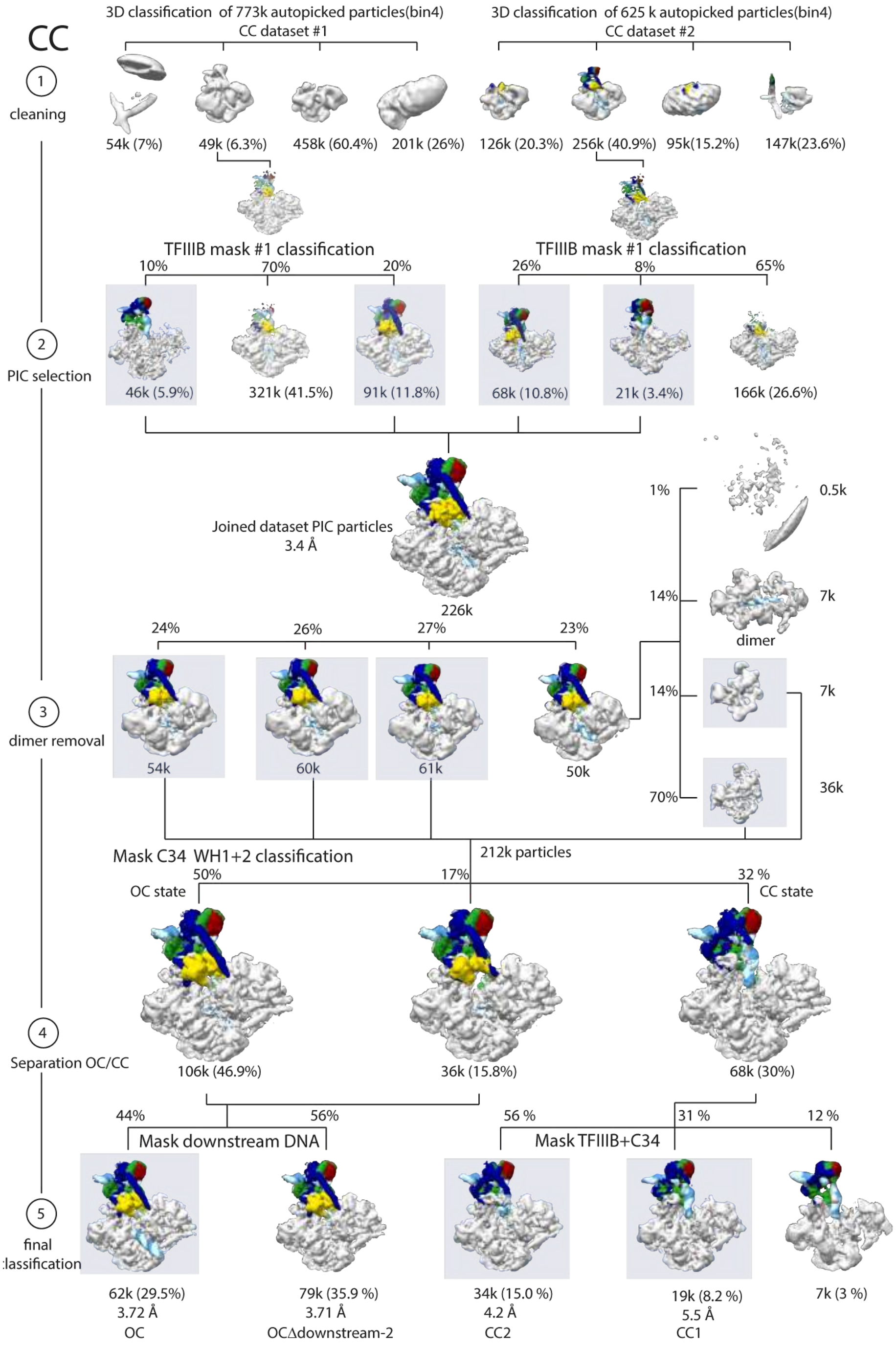

CC
3D classification of 773k autopicked particles(bin4)
CC dataset #1
3D classification of 625 k autopicked particles(bin4)
CC dataset #2
1
cleaning
54k (7%)
49k (6.3%)
458k (60.4%)
201k (26%)
126k (20.3%)
256k (40.9%)
95k(15.2%)
147k(23.6%)
TFIIIB mask #1 classification
10%
70%
20%
TFIIIB mask #1 classification
26%
8%
65%
2
PIC selection
46k (5.9%)
321k (41.5%)
91k (11.8%)
68k (10.8%)
21k (3.4%)
166k (26.6%)
Joined dataset PIC particles
3.4 Å
226k
1%
0.5k
14%
7k
dimer
14%
7k
70%
36k
24%
26%
27%
23%
3
dimer removal
54k
60k
61k
50k
Mask C34 WH1+2 classification
212k particles
50%
OC state
17%
32 %
CC state
4
Separation OC/CC
106k (46.9%)
36k (15.8%)
68k (30%)
44%
56%
Mask downstream DNA
56 %
31 %
12 %
Mask TFIIIB+C34
5
final
classification
62k (29.5%)
3.72 Å
OC
79k (35.9 %)
3.71 Å
OCΔdownstream-2
34k (15.0 %)
4.2 Å
CC2
19k (8.2 %)
5.5 Å
CC1
7k (3 %)

Figure 2.5: Classification of particles from the CC datset. Two CC datasets were individually cleaned (step 1) by classification in 3D (binned 4 times) and PIC particles were selected using a mask on TFIIIB (step 2). PIC particles from both datasets were combined and classified without masking to remove a small subset of dimers (step 3). Classification using a mask on C34 WH1 and WH2 yielded the OC and CC populations (step 4) which were subsequently classified into CC1, CC2, OC and OCΔdownstream-2 (step 5). The majority of particles from the CC datasets are in the OC state (67% of PIC particles vs 32% in the CC states, 2.5), showing that the TFIIIB-Pol III complex is active in promoter opening. Figure taken from [203].

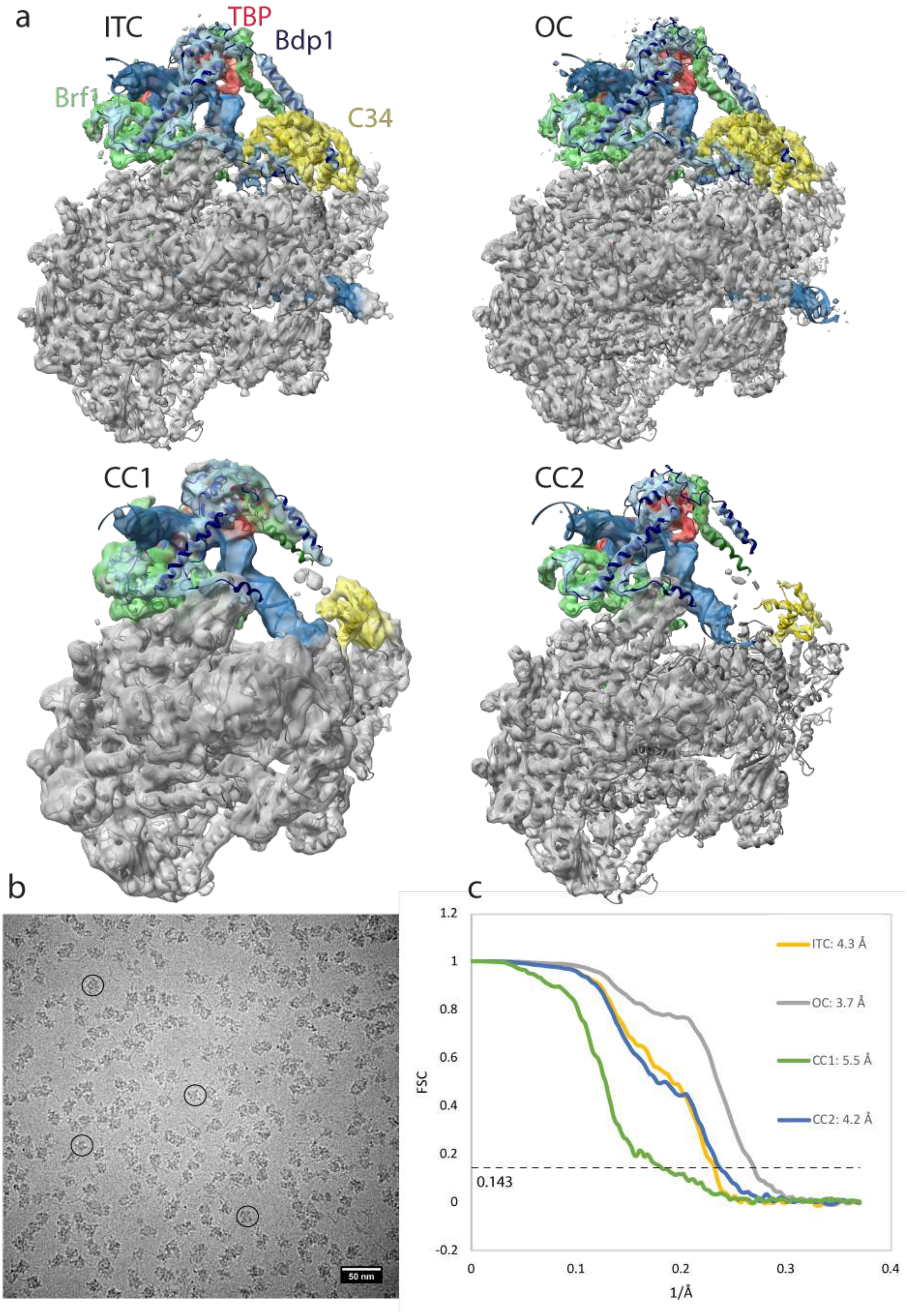


Figure 2.6: a. Cryo-EM densities of the Pol III ITC, OC, CC1 and CC2 maps superimposed with the corresponding models. b, Typical high defocus micrograph of the Pol III PIC. c, Fourier-shell correlation (FSC) curves for the different cryo-EM maps as reported by the relion postprocess program. Figure taken from [203].

during classification (see previous paragraph). In the ITC reconstruction, DNA density is well-defined and RNA is visible at a lower threshold. This suggests partial occupancy of the RNA, owing either to dissociation during sample preparation or cleavage by C11.

During particle classification of the CC dataset, it became apparent that the PIC had spontaneously opened the closed DNA in the majority of particles, giving rise to a well resolved OC at 3.7 Å. The remaining particles from the CC dataset were found in two different conformations, termed CC1 and CC2.

The OC was determined at 3.7 Å resolution and is almost identical to the ITC, but shows only fragmented density for the template strand in the active site.

The two CC reconstructions (CC1 and CC2, at 5.5 Å and 4.2 Å resolution, respectively) differ in the orientation of TFIIIB with respect to Pol III, and show a well-defined Pol III core with markedly lower resolution at the periphery (Figure 2.7). I built atomic models of the PIC in the CC, OC and ITC states. Local amplitude scaling (LocScale) [88] of the maps supported model building. The TFIIIB structure comprises most of Brf1 , whereas for Bdp1 it includes ER II (residues 275–320) and ER I (residues 365–537) (Figure 2.9). In addition, the WH1 and WH2 domains of C34, and a previously disordered ‘initiation/termination loop’ (residues 211–224) in subunit C37, which has been shown to have important roles in both transcription initiation and termination [85, 107, 164, 214], become ordered in the ITC and OC structures. The maps are depicted in Figure 2.6.

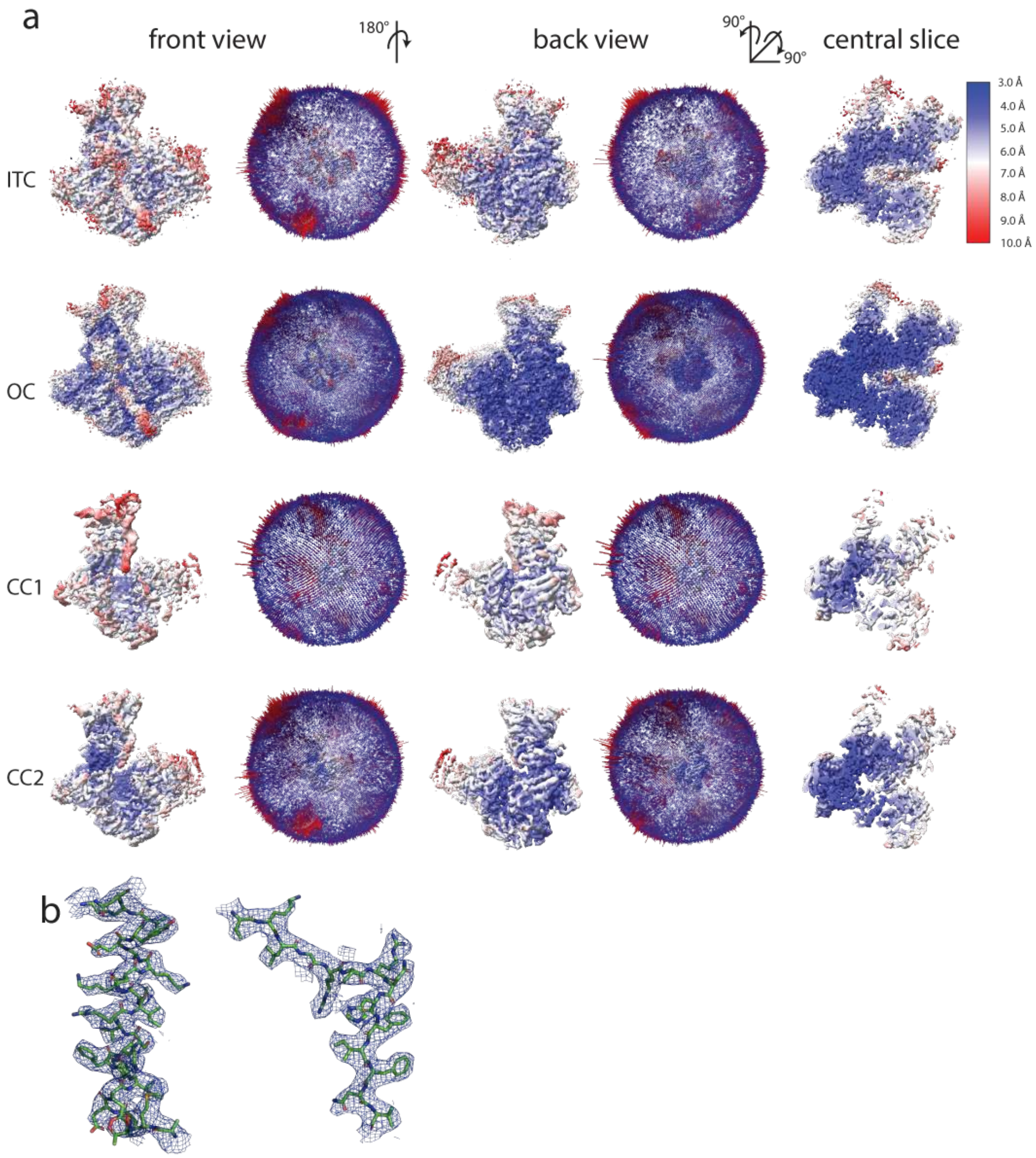

a
front view
180°
back view
90°
90°
central slice
3.0 Å
4.0 Å
5.0 Å
6.0 Å
7.0 Å
8.0 Å
9.0 Å
10.0 Å
ITC
OC
CC1
CC2
b

Figure 2.7: Cryo-EM maps colored by local resolution from 3 Å (dark blue) to 10 Å (red). All maps are shown on the same contour level. Central slices show that the local resolution degrades more strongly in the CC2 map compared to the ITC map, although they have comparable overall resolutions. Euler angular distribution plotting shows a good angular coverage without a dominating preferred orientation in all reconstructions. b, Examples of helical and β-strand densities in Pol III subunit AC40 in the 3.7 Å OC map. Figure taken from [203].

2.2.4 *Architecture of the Pol III open complex / initially transcribing complex*

The cryo-EM maps allowed me to built the most complete models of TFIIIB to date. Figure 2.9 shows these structures in isolation, and Figure 2.8 in context of the PIC and might help to guide the reader through the following sections.

In the Pol III PIC, TFIIIB is located in a similar position as TFIIB-TBP are in the Pol II PIC. The N-terminal Zn-ribbon is located near the active site and seals the RNA exit tunnel (Figure 2.8). A primer extension assay with the ITC scaffold in absence or presence of TFIIIB revealed an increase in transcription pausing at nt 11 in presence of TFIIIB, indicating that the RNA chain clashes with the Brf1 Zn-ribbon when it reaches 12 nt, and presumably this is the critical length of the nascent RNA required for promoter esacpe (Figure 2.10).

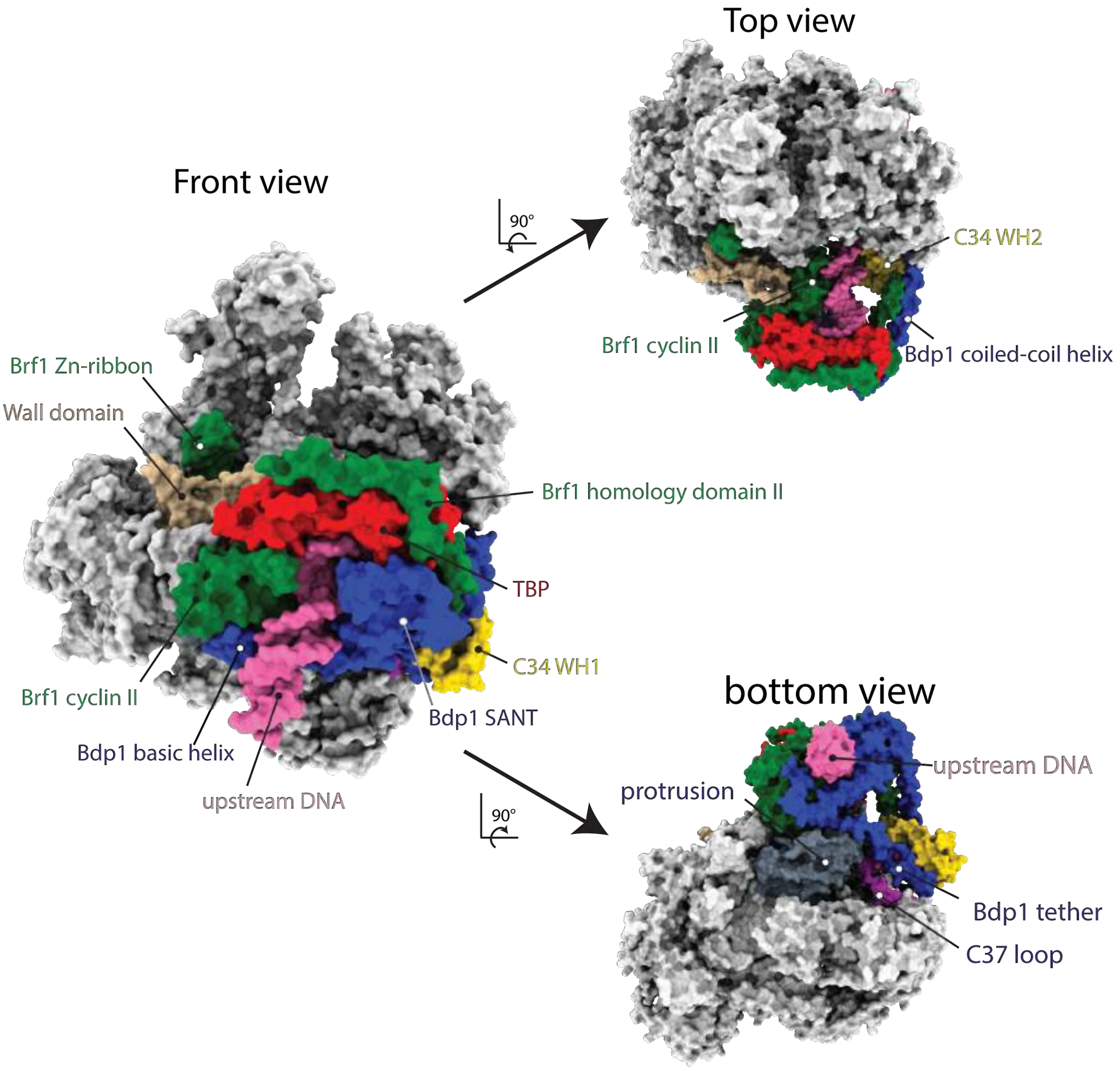

Top view
Front view
90°
C34 WH2
Brf1 cyclin II
Bdp1 coiled-coil helix
Brf1 Zn-ribbon
Wall domain
Brf1 homology domain II
TBP
C34 WH1
Bdp1 SANT
Brf1 cyclin II
Bdp1 basic helix
upstream DNA
bottom view
90°
protrusion
upstream DNA
Bdp1 tether
C37 loop

Figure 2.8: Overview of the open complex Pol III-PIC . Important structural elements and domains in Pol III are labelled. Based on PDB 6f40.

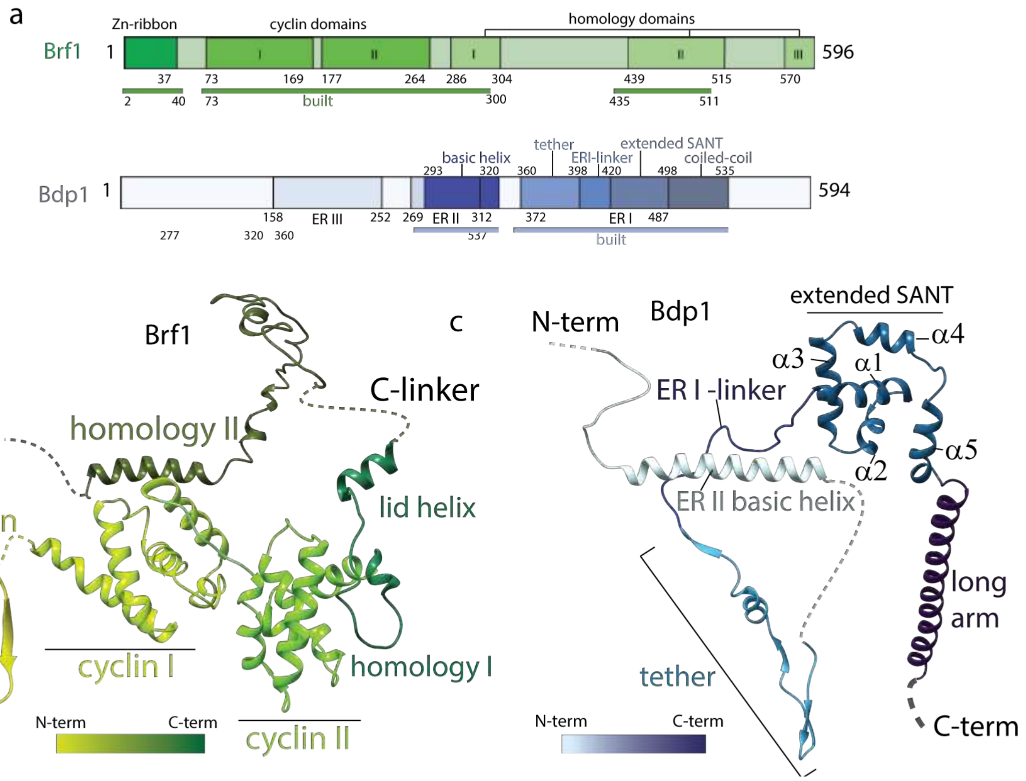

a
Zn-ribbon
cyclin domains
homology domains
Brf1
1
I
II
I
II
III
596
37
73
169
177
264
286
304
439
515
570
2
40
73
built
300
435
511
tether
extended SANT
basic helix
ERI-linker
coiled-coil
293
320
360
398
420
498
535
Bdp1
1
594
158
ER III
252
269
ER II
312
372
ER I
487
277
320
360
537
built
b
Brf1
C-term
homology II
C-linker
N-ribbon
lid helix
cyclin I
homology I
N-term
C-term
cyclin II
c
N-term
Bdp1
extended SANT
α3
α1
α4
ER I -linker
α5
α2
ER II basic helix
long arm
tether
N-term
C-term
C-term

Figure 2.9: Strucutres of Brf1 and Bdp1 in the PIC. a: Annotated domain structure of Brf1 and Bdp1. Parts of TFIIIB that are included in the ITC model are underlined and marked as 'built'. b: ribbon representation of Brf1, coloured in shades of green from N- to C-terminus. c: ribbon representation of Bdp1, coloured in shades of blue from N- to C-terminus.

The linker connecting the Zn-ribbon with Brf1 cyclin I is disordered. The Brf1 homology domain I, which follows cyclin II, forms a short 'lid' helix that lies on top of the DNA (Figure 2.11). This helix is equivalent to the redox sensing 'molecular pin' in human Brf2 [65], but lacks oxidizable residues. The residues connecting homology domains I and II are disordered, but weak density in proximity to the Bdp1 SANT domain is visible at a high threshold which is tentatively assigned to this region, in agreement with photo-crosslinking and mutagenesis studies [116].

The residues of homology domain II also show relatively weak density, but could be modelled by fitting a crystal structure comprising this region (PDB 1ngm). Homology domain II wraps around TBP and leads into a helix that runs in parallel to promoter DNA, but without contacting it (Figure 2.8, panel "Top view"). I could not model any residues belonging to Brf1 homology domain III, although I observed weak density close to C34 WH3; this may correspond to Brf1, because Brf1 residue 549 photo-crosslinks to C34 [116]. The cyclin domains I and II bind the Pol III wall and upstream promoter DNA (Figure 2.8). They are part of a compact core around upstream promoter DNA, in which the TATA box and adjacent bases are held in place by a positively charged ring that completely surrounds the DNA (Figures 2.8, 2.11).

Bdp1 contributes to this core via its most conserved elements. The most N-terminal, ordered part of Bdp1 corresponds to ER II, in which residues 277–292 interact with cyclin II of Brf1, and residues 293–319 form a basic helix containing 14 positively charged residues

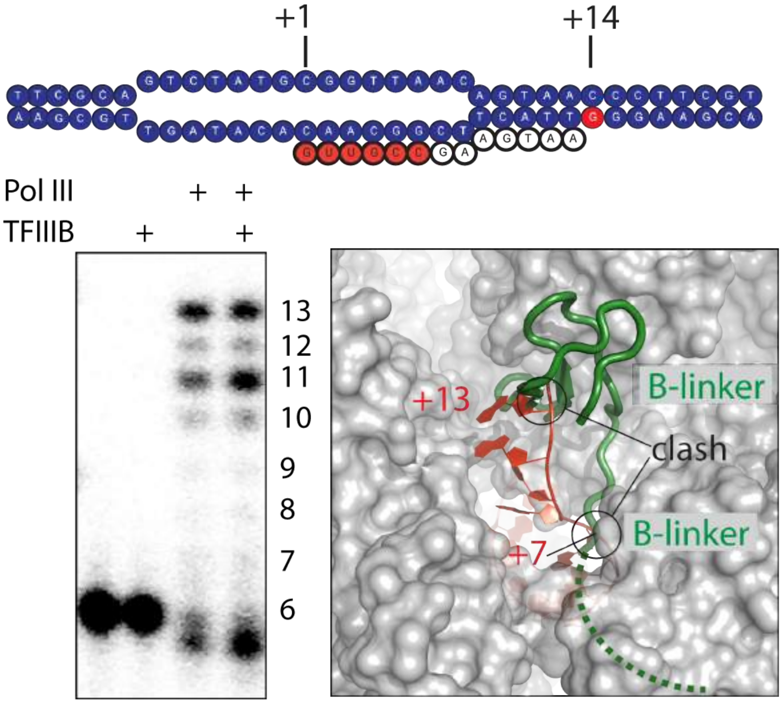

+1
+14
TTCGCA GTCTATGCGGTTAAC AGTAACCCTTCGT
AAGCGT TGATACACAACGGCT TCATTGGGAAGCA
GUUGCC GA AGTAA
Pol III + +
TFIIIB + +
13
12
11
10
9
8
7
6
+13
+7
B-linker
clash
B-linker

Figure 2.10: RNA extension assay of the ^{32}P-labelled 6-nt RNA in absence of CTP, showing that the preparation is active in transcription using the ITC scaffold. Addition of TFIIIB leads to a stronger accumulation of 11-nt RNA. Modelling of an elongated RNA oligonucleotide based on Pol I (PDB 5m5x) shows that RNA clashes at positions +7 with the beginning of the flexible B-linker, and at position +13 with the B-ribbon. Accumulation of 11-nt RNA in presence of TFIIIB suggests that the RNA in the Pol III-PIC takes a different path compared to the Pol I elongation complex, and clashes with the B-ribbon at position +12, requiring release of TFIIIB to enter elongation. Primer extension assay performed by Florence Baudin (F.B.).

that runs behind upstream DNA (Figure 2.11). The residues connecting ER II to ER I are disordered. ER I is located between C34 WH2 and the heterodimer, forming a β-hairpin that interacts with WH1 and WH2 of C34, and with the C37 initiation/termination loop, in agreement with photo-crosslinking data [85, 214]. Bdp1 then forms a short helix that runs along the protrusion domain of Pol III, followed by a β-strand that extends the anti-parallel β-sheet of the Pol III protrusion by one parallel strand. This section of Bdp1 (residues 360–398)—comprising the β-hairpin, the helix and the β-strand—interacts with three distinct Pol III elements: the protrusion, the previously disordered WH1 and WH2 of C34, and the C37 initiation/termination loop. I named this element the Bdp1 tether, owing to its central position in the PIC that bridges between distant and/or mobile Pol III subunits. The tether connects back to the DNA binding TFIIIB core, through the ER I linker, which is enriched in aromatic residues and inserts into the minor groove of the DNA. This linker leads into an extended SANT domain, which is located on the opposite site of the Brf1 cyclin domains (Figure 2.11). The canonical SANT domain fold is extended by two C-terminal helices (Figures 2.8, 2.9) and is followed by a long helix (residues 498–536) that forms a coiled-coil with Brf1 homology domain II and ends next to C34 WH2 and WH3. In summary, TFIIIB forms a compact core centred around the TATA box, from which two arms emerge. A short arm, formed by the Bdp1 tether, and a long arm, formed by the Bdp1–Brf1 coiled-coil, bridge the TFIIIB core with the Pol III heterotrimer and heterodimer.

2.2.5 *DNA binding mode of TFIIIB explains long residence time on chromatin*

Our cryo-EM structure reveals DNA-binding elements absent from previous crystallographic studies of TFIIIB [66, 95] and Pol II paralogues, namely the Brf1 lid helix and the Bdp1 basic helix. Together with the Brf1 cyclin folds, TBP and the Bdp1 SANT and linker domains, they form a positively charged ring around the TATA box (Figures 2.11, 2.8), explaining why promoter bound TFIIIB is so unusually stable [99, 101, 103, 120]. All three subunits are required to completely surround the DNA, and this 'headlock' arrangement illustrates how Bdp1 causes the kinetic trapping of the TFIIIB–DNA complex [34]. Accordingly, the term 'proteinaceous cage', which was coined to describe the TFIIIB–DNA structure almost two decades ago (cited in [34]), is an excellent description of this structure.

The binding interface is rich in positively charged residues. This suggests that the DNA component is essential for the stability of the entire complex, and explains why TFIIIB dissociates into its constituents B′ (Brf1-TBP) and B″ (Bdp1) during chromatographic separation.

As pointed out in section 1.3.2, the *in vitro* life-time of the TFIIIB-DNA complex exceeds the doubling time of log-phase yeast, and *in vivo* occupancy studies show that TFIIIB stays bound to Pol III promoters under repressive conditions. This provides an opportunity for rapid re-activation, and resembles the situation of poised Pol II genes, which contain promoter-proximally stalled Pol II complexes to allow the rapid transcription of fast responding genes, such as the heat shock chaperone Hsp70 [170].

However, the long lifetime of TFIIIB also affects other nuclear processes, as it can serve as a 'roadblock' for Pol II transcription and DNA replication by DNA polymerase [172].

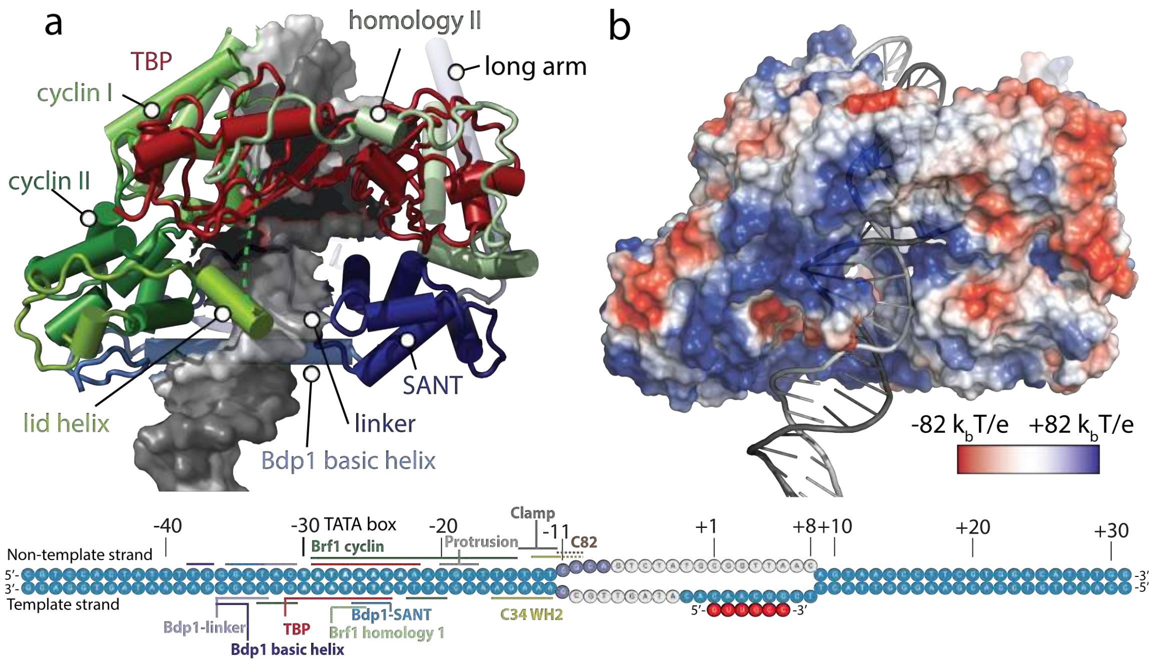

a
homology II
TBP
long arm
cyclin I
cyclin II
SANT
lid helix
linker
Bdp1 basic helix
b
-82 k_bT/e
+82 k_bT/e
Clamp
-40
-30 TATA box
-20
-11
+1
+8+10
+20
+30
Non-template strand
Brf1 cyclin
Protrusion
C82
5'-
3'-
-3'
-5'
Template strand
Bdp1-linker
TBP
Bdp1-SANT
Brf1 homology 1
C34 WH2
Bdp1 basic helix
5'-
-3'

Figure 2.11: TFIIIB forms a positively charged ring around upstream promoter DNA. a: cylinder representation of TFIIIB around the TATA box with DNA represented as surface. b: electrostatic potential mapped onto the TFIIIB structure. Bottom: representation of the DNA scaffold used for the ITC dataset. Protein-DNA contacts are indicated. Grey circles represent bases in the transcription bubble that were disordered. Adapted from [203].

2.2.6 *TFIIIB coordinates promoter opening*

2.2.6.1 *Allosteric changes in Pol III*

TFIIIB triggers conformational changes in Pol III that are specific to the preinitiation state and/or render Pol III ready for elongation. In the PIC, Pol III adopts a conformation similar to that of the closed clamp state seen in transcribing Pol III [82], however the heterotrimer shifts further towards upstream DNA, closing the clamp slightly more, and the stalk moves towards the heterotrimer, possibly mediated by a density that is visible between the stalk and heterotrimer which is assigned to the C31 subunit [82], although the quality of the density is insufficient for modelling.

The C53–C37 heterodimer contributes to the stabilization of the PIC by (partial) ordering of the initiation/termination loop, which interacts with the Bdp1 tether and inserts between WH1 and WH2 of C34, concomitant with the extension of the adjacent C37 helix (residues 197–202), but shows no large scale conformational changes.

The C-terminal domain of C11 is disordered, resembling the elongating state. In the active site, the rudder and the trigger loop are disordered and the bridge helix is bent, as in elongating Pol III (Fig 2.12).

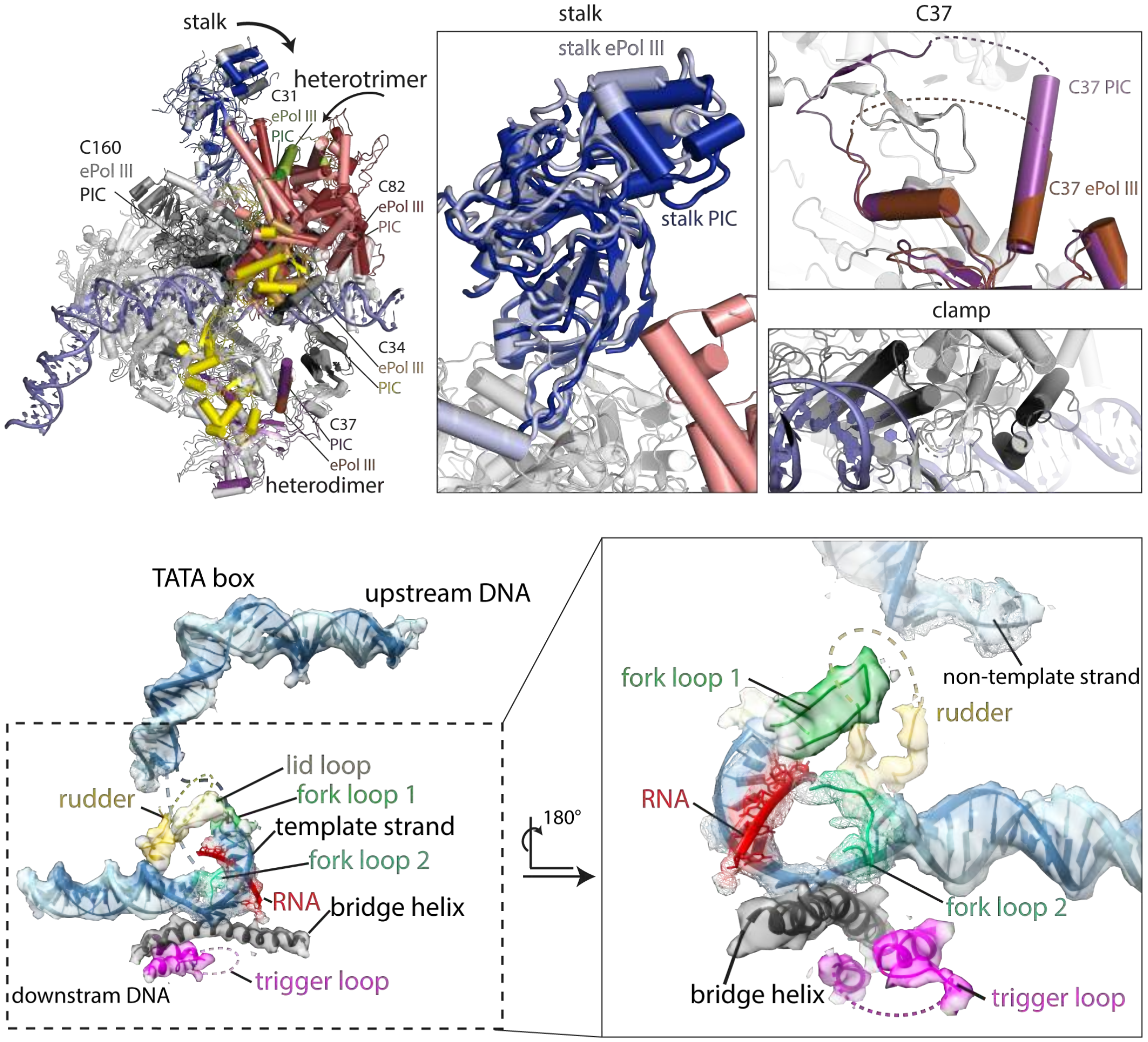

stalk
heterotrimer
C31
ePol III
PIC
C160
ePol III
PIC
C82
ePol III
PIC
C34
ePol III
PIC
C37
PIC
ePol III
heterodimer
stalk
stalk ePol III
stalk PIC
C37
C37 PIC
C37 ePol III
clamp
TATA box
upstream DNA
rudder
lid loop
fork loop 1
template strand
fork loop 2
RNA
bridge helix
trigger loop
downstream DNA
180°
fork loop 1
non-template strand
rudder
RNA
fork loop 2
bridge helix
trigger loop

Figure 2.12: Top: Superposition of elongating Pol III (ePol, PDB 5fj8) and Pol III in the OC. In the OC, the heterotrimer moves towards upstream DNA and the C34 WH1 and WH2 domains become ordered. The stalk moves towards the heterotrimer, and the clamp moves to slightly close the cleft. The C37 initiation/termination loop becomes partially ordered to interact with the Bdp1 tether and C34 WH1 and WH2. EM density is shown in the ITC map after amplitude scaling of active site elements and nucleic acids. Elements contoured at a lower threshold are shown as mesh. Active site elements are labelled. The rudder and trigger loops are disordered. Poor density for RNA suggests flexibility or partial occupancy due to dissociation during sample preparation or cleavage by C11. Figure adapted from [203].

2.2.7 *The transcription bubble is stabilized by Pol III, but coordinated by TFIIIB*

Melting of the DNA duplex and stabilization of the resulting transcription bubble is arguably the central step in transcription initiation. In the reconstruction obtained in the initially transcribing state, I observed opening of the DNA duplex at position -11 relative to the TSS, although the DNA scaffold that was used was fully complementary until position -8, showing that the PIC had extended the transcription bubble. Density supported modelling of four nucleotides of the unpaired non-template strand, and of ten nucleotides of the template strand in the active site that were base-paired to the

RNA. The remainder of the single-stranded DNA was not observed, indicating disorder. Interestingly, no density for single-stranded DNA was observed in the spontaneously formed open DNA complex, which presumably results from a heterogenous size of the transcription bubble formed from fully complementary DNA.

The upstream edge of the transcription bubble and adjacent duplex DNA (bp -20 to -10) are stabilized by four different elements: the C34 WH2, the C82 WH4 including the cleft loop, the clamp core of C160, and the Brf1 cyclin I (Figures 2.11, 2.13). Notably, most of these elements are Pol III subunits, with few direct DNA contacts of TFIIIB. However, the C34 WH2 domain—which is disordered in structures of unbound and transcribing Pol III—is positioned by a network of interactions with TFIIIB, including the Bdp1 tether and the long arm, and the initiation/termination loop of C37 (Figure 2.14). Mutations in the C34 WH domains compromise promoter opening [21, 205], and the presented structures show C34 WH2 binding to the melted non-template strand, corroborating its essential function.

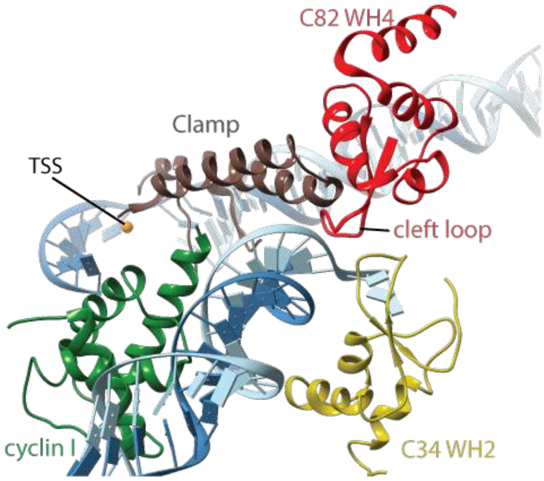


Figure 2.13: DNA binding around the transcription bubble. The transcription bubble is stabilized by Brf1 cyclin I, the Pol III clamp, C34 WH2 and C82 WH4. The non-template DNA strand is stabilized by the C82 cleft loop and C34 WH2. Figure taken from [203].

Furthermore, deletion of the Bdp1 tether, which forms the C34 binding platform, specifically prevents melting of the upstream edge of the transcription bubble [100].

The spatial organization of the PIC, with the Brf1 Zn-ribbon and N-linker close to the active site, also rationalizes why a Brf1Δ(1-68) mutant requires pre-opening of the DNA around the TSS for transcriptional activity. Moreover, the lack of direct DNA contacts between TFIIIB and DNA around the transcription bubble agrees with the unchanged DNA-crosslinkining pattern in open and closed DNA complexes (see section 1.3.4).

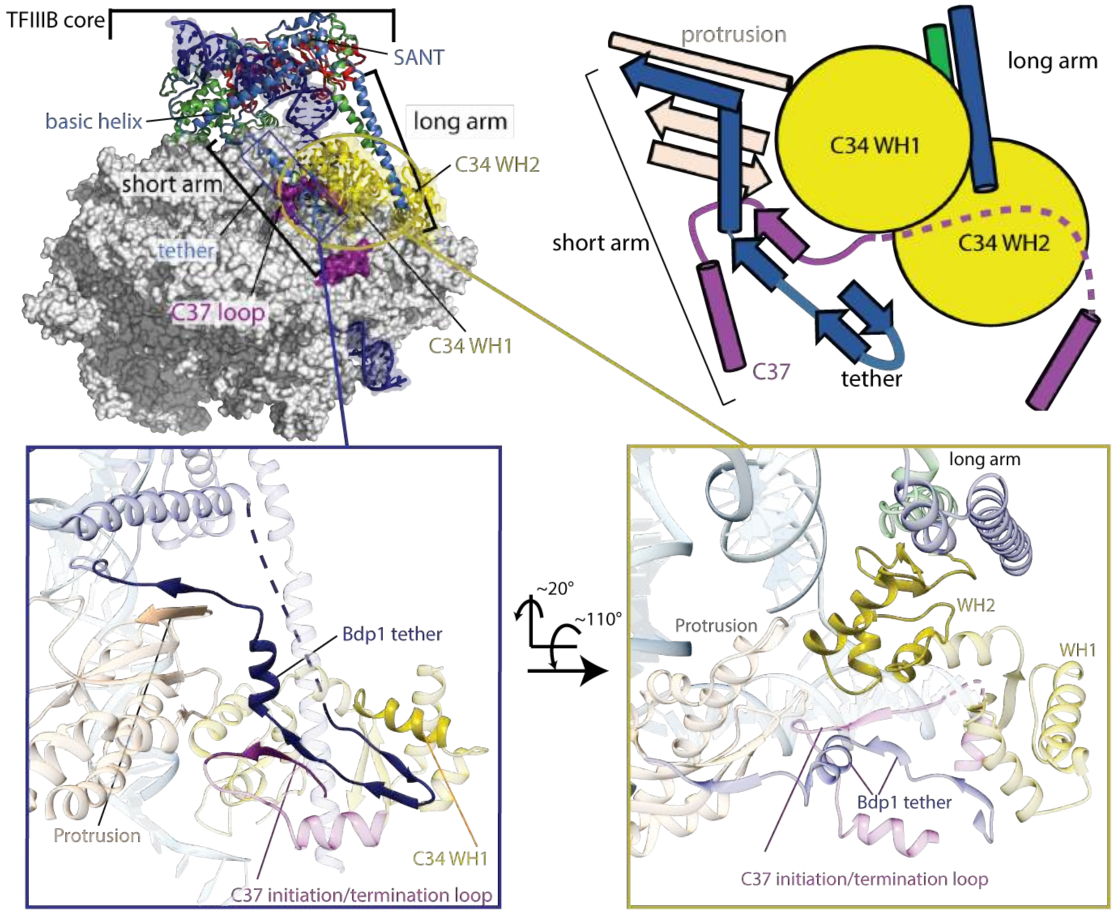

TFIIIB core
SANT
basic helix
long arm
C34 WH2
short arm
tether
C37 loop
C34 WH1
protrusion
long arm
C34 WH1
C34 WH2
short arm
C37
tether
Bdp1 tether
Protrusion
C34 WH1
C37 initiation/termination loop
~20°
~110°
long arm
Protrusion
WH2
WH1
Bdp1 tether
C37 initiation/termination loop

Figure 2.14: Top panel: Location of TFIIIB on Pol III (left) and schematic view of interactions that stabilize C34 in the position observed in the OC/ITC structures (right). Bottom left: Close up view of the Bdp1 tether. Interactions between Bdp1 tether, Pol III protrusion, C37 initiation/termination loop and C34 WH1 and WH2 are highlighted. Regions that contact the Bdp1 tether are depicted in solid color, other elements are transparent. Bottom right: Close up view of C34 WH2. C34 WH2 is stabilized by both arms of TFIIIB, Pol III protrusion and C37. Figure taken from [203].

2.2.8 *Structures of the closed DNA complexes*

The Pol III closed DNA structures (Figure 2.15) show early engagement intermediates of the PIC, as several of the stabilizing interactions seen in the OC and ITC are not yet established. The C34 WH1 and WH2 are disordered, as are the Bdp1 tether and most of the long arm of TFIIIB and the C37 initiation/termination loop. This suggests that C34 and the TFIIIB arms mutually stabilize each other, and that these interactions can occur only once the DNA has been melted or moved to slightly enter the cleft (discussed below). Accordingly, the resolution in our CC maps declines more strongly in the peripheral regions around the closed DNA and the heterotrimer compared to the OC and ITC maps (Figure 2.15). The upstream DNA is kinked away by the clamp head and the C82 cleft loop in the CCs, resulting in a ~30° bend introduced around position -15. The DNA projects away from Pol III, and no density is visible after bp -2 (bp -5 in CC2), presumably because of the lack of stabilizing protein–DNA contacts (Figure 2.15). The CC1 and CC2 reconstructions differ in the orientations of TFIIIB, upstream DNA and the heterotrimer.

In CC1, the heterotrimer is moved away from upstream DNA compared to its orientation in CC2, which slightly opens the clamp (Figure 2.15). In addition, TFIIIB is shifted by up to 6 Å, which moves the DNA away from the cleft (Figure 2.15), whereas CC2 exhibits an overall conformation similar to that seen in the OC and ITC states. This suggests that CC2 represents a later stage in the initiation process. The cryo-EM density of DNA extends further in the CC1 map, whereas that of TFIIIB is of better quality in the CC2 map.

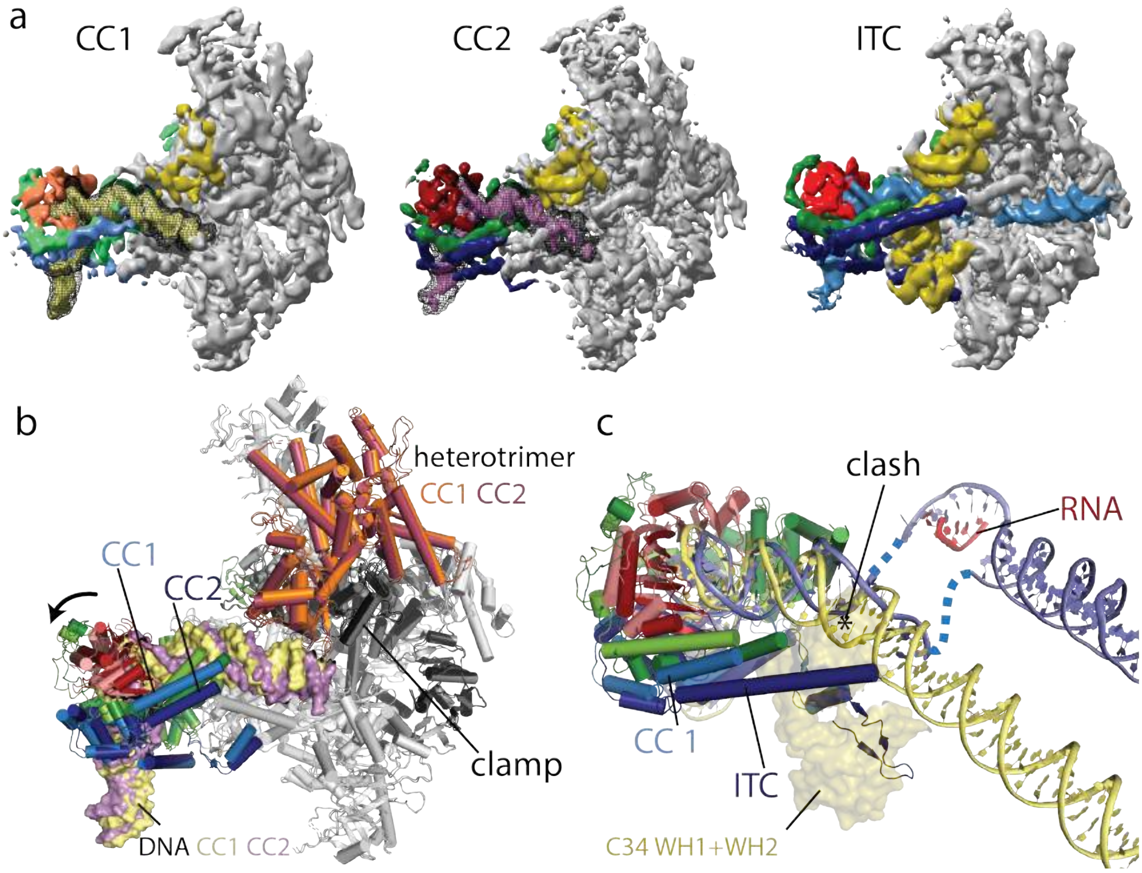


Figure 2.15: Comparison between Pol III closed complex and open complex structures. a, EM densities of the CC1, CC2 and ITC with densities corresponding to TBP colored in orange/red, Brf1 colored in green, Bdp1 colored in blue, C34 colored in yellow. CC1 is colored in lighter shades and DNA is colored in pale yellow for CC1 and magenta for CC2. DNA density before B-factor sharpening is shown in addition as black mesh for CC1 and CC2. In the CCs, C34 WH1 and WH2 and the C-terminal part of the Bdp1 coiled-coil helix are disordered. b, Superposition of CC1 and CC2 models shows movement of TFIIIB and the heterotrimer, slightly closing the cleft during the CC1 to CC2 transition. c, Superposition of CC1 and ITC models. DNA for CC1 was further extended with B-DNA to help visualization. The positions of C34 WH1 and WH2 (transparent surface) in the ITC would clash with CC DNA. All models were superimposed on C160. Figure taken from [203].

2.2.9 *Mechanistic model of promoter opening*

I have combined the structures reported here to obtain a model of the DNA melting process (Figure 2.16, and Supplementary Video 1 of ref [203]). I also included a modelled intermediate (CC, open clamp), as I noticed that our CC structure is markedly different from those of yeast and human Pol II. In the Pol II system, DNA runs along the length of the polymerase and interacts with the jaw domain at the downstream end [77, 159]. In our structure, DNA can be traced only until bp -2, and projects away from Pol III. The clamp is in a closed state, the C82 cleft loop blocks access to the cleft, and the C34 WH1 and WH2 are disordered, presumably because they cannot occupy the same position as they do in the OC because they would clash with DNA (Figure 2.15). It is therefore likely that the transition from the CC to the OC proceeds via an intermediate in which the clamp is open, similar to human Pol II and bacterial RNA polymerase [54, 77]. I therefore modelled a CC, open-clamp intermediate based on the structure of apo Pol III [82] .

This model was obtained by combining the structure of open clamp apo Pol III (PDB: 5FJA) with the DNA of the human Pol II PIC in its closed state (PDB: 5iya); this DNA has a much stronger resemblance to that of the Pol III OC than does DNA of the yeast Pol II OC. The position of C34 was obtained by superimposing C34 from the Pol III OC model onto the open-clamp structure and aligning the WH3 domains. In this model there are only minor clashes between DNA and the Rpb5 subunit in the jaw, which could easily be accommodated by a slightly different curvature of the DNA.

In the open clamp, closed complex model, the clamp and the cleft loop move upwards and lie on top of the closed DNA. The C34 winged helices can adopt similar positions as they do in the OC because the steric clash with DNA is removed, but they are in closer proximity to the DNA compared to their positions in the OC. This would trap DNA in the clamp, and clamp closing would enforce DNA melting by a steric clash of the C82 cleft loop and the DNA duplex, loading the template strand into the active site. This process could be driven by the stabilization of the C34 WH2 in the OC conformation by Bdp1, C37 and the protrusion, as described earlier.

This model is similar to what has been put forward for Pol II [77, 159], as both use the coupled movements of the clamp, an extended loop (E-wing/C82 cleft loop) and winged-helix domains (heterotrimer/TFIIE–TFIIF) for promoter opening and nascent bubble stabilization. This suggests that the basic mechanism of promoter opening is conserved between Pol II and Pol III.

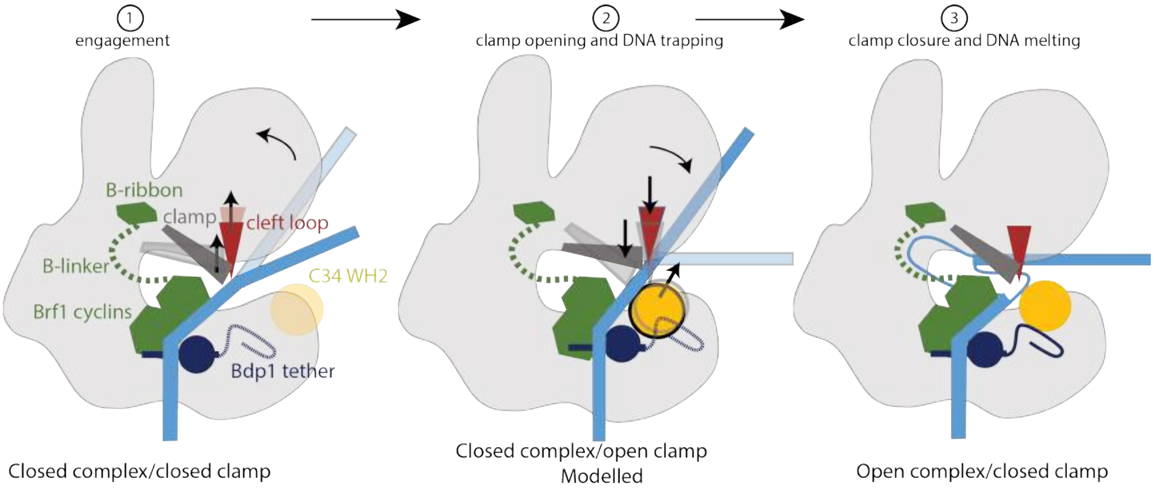

①
engagement
②
clamp opening and DNA trapping
③
clamp closure and DNA melting
B-ribbon
clamp
cleft loop
B-linker
C34 WH2
Brf1 cyclins
Bdp1 tether
Closed complex/closed clamp
Closed complex/open clamp
Modelled
Open complex/closed clamp

Figure 2.16: Schematic drawing of the promoter opening. Initially, the closed DNA is bent away by the C82 cleft loop and the clamp, while the Bdp1 tether and C34 WH1 and WH2 are disordered. Opening of the clamp allows the closed DNA to slide between clamp and lobe. Subsequently, C34 WH1 and WH2 become ordered and enclose double-stranded DNA. Closing of the clamp leads to the downwards movement of the cleft loop and DNA melting, while establishing stabilizing interactions between the Bdp1 tether and C34 WH1 and WH2. The non-template strand is stabilized by Pol III subunits C34 and C82; the template strand is presumably stabilized by the Brf1 linker. Figure taken from [203].

3

MOLECULAR BASIS OF POL III REPRESSION BY MAF1

3.1 RECONSTITUTION OF THE MAF1-POL III COMPLEX AND INITIAL EM RECONSTRUCTIONS

The Maf1-Pol III complex was reconstituted in absence of nucleic acids and binding was tested by size exclusion chromatography. Formation of a stable complex required low ionic strength, but never appeared stochiometric judged by SDS-PAGE analysis. Efforts to improve binding of Maf1 included testing of N- and C-terminally tagged Maf1 constructs, and Pol III purified from strains carrying affinity tags on either the AC40 subunit or the C128 subunit. Initial analysis suggested that the C128 tagged strain resulted in improved binding, therefore subsequent experiments were performed with this strain. However, during the course of this project it became apparent that a large batch-to-batch variability exists even for identical strains. I suspect that the underlying cause for this lies in the growth phase of the yeast culture at the point of harvest, where batches grown to a slightly higher optical density might accumulate repressive post-translational marks on Pol III, which favours binding of Maf1.

Under low salt conditions, full-length Maf1 had the tendency to precipitate. This was also evident from cryo-EM samples prepared in 40 mM $(NH_4)_2SO_4$ with Pol III or mixtures of Pol III and Maf1 (Figure 3.1).

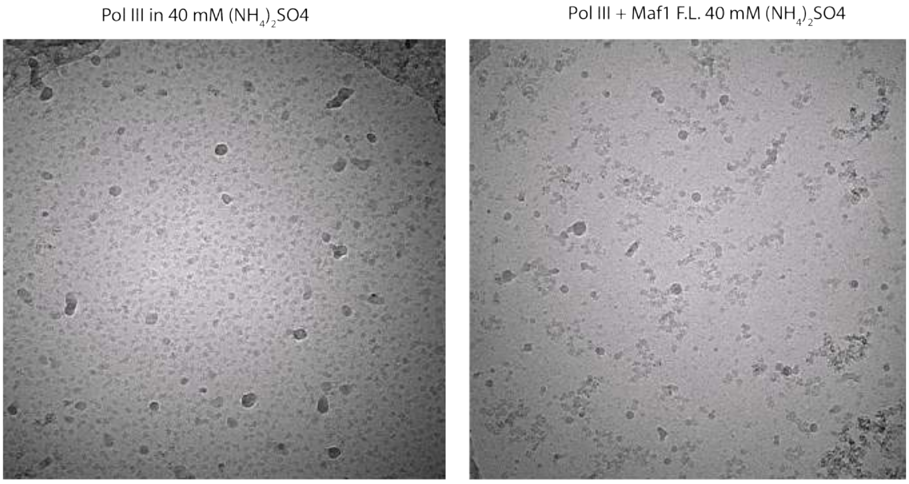

Pol III in 40 mM $(NH_4)_2SO4$
Pol III + Maf1 F.L. 40 mM $(NH_4)_2SO4$

Figure 3.1: Full-length Maf1 aggregates under low-salt conditions. Cryo micrographs of Pol III (left) or Pol III incubated with full-length Maf1 (right) show that full-length Maf1 causes aggregation under low-salt conditions.

I therefore resorted to a construct where a non-essential, non-conserved stretch of disordered residues is deleted, which is functional *in vivo* [142]. This construct is referred to as Maf1 i.d. (for internal deletion).

A small initial cryo-EM dataset was aquired on a Titan Krios. For this dataset, Maf1 i.d. was mixed with Pol III and applied to holey Mo 1.2/1.3 grids without further purification. In this sample, I observed strong dissociation of Pol III subcomplexes, namely the heterotrimer and the heterodimer (Figure 3.2). In addition, the resolution of the 3D reconstructions was highly anisotropic, indicating a strongly preferred orientation. However, classification revealed a subset of particles with very weak additional density close to the Pol III wall. After further classification which was focused on this density (see Figure 3.2), I could rigid body fit the crystal structure of human Maf1 into this density, which allowed me to use model-based local amplitude scaling as implemented in the program LocSclae [88], which greatly improved map interpretability. The central β-sheet of Maf1 was visible as a flat sheet of density, and some helical density was also observed. This density fitted the structure of human Maf1 reasonably well, but the low occupancy of Maf1, as well as the very anisotropic resolution of the 3D reconstruction, necessitated further sample optimization. It is worth pointing out that the approximate location of Maf1 in this map disagreed with the published model [201], which motivated me to obtain high-resolution information.

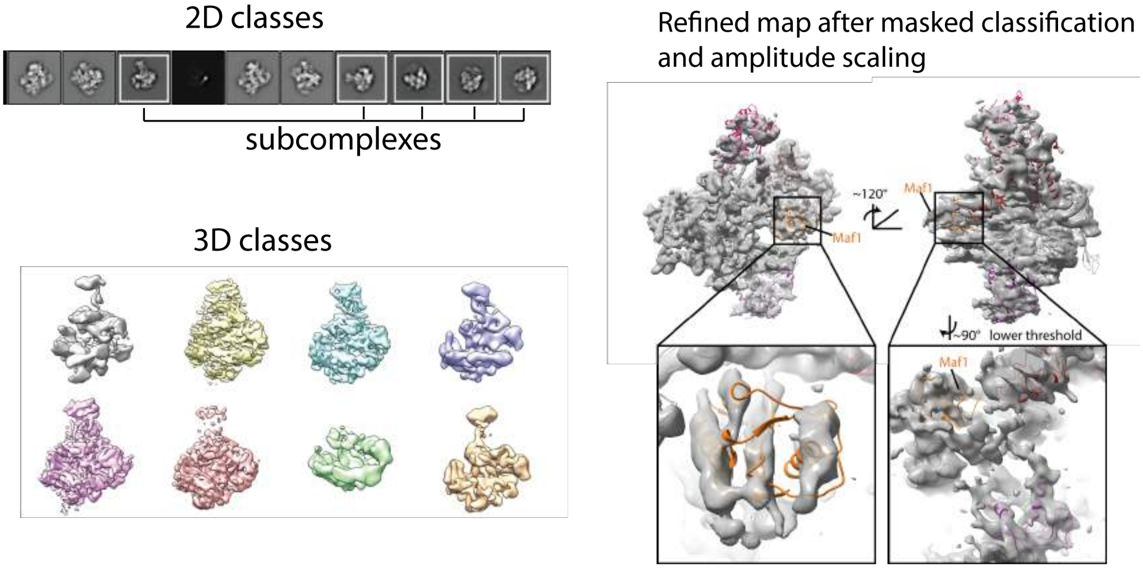


Figure 3.2: Initital cryo-EM reconstruction of Maf1-Pol III. Using Mo 1.2/1.3 girds, strong dissociation is visible in 2D and 3D classes. A map obtained after focussed classification and local amplitude scaling restores density that can be fitted with the Maf1 crystal structure.

At the same time, I decided to pursue crystallization of the yeast Maf1 i.d. construct in order to obtain a reliable model for density fitting, in case no high resolution map could be obtained.

3.2 CRYSTALLIZATION OF MAF1

The initial construct of the internal deletion Maf1 i.d. (provided by Ian Willis from the Einstein College, New York) contained a non-cleavable C-terminal his tag and failed to crystallize, although the protein could be expressed in large amounts and was highly homogeneous. To increase the chances of crystallization, I cloned a construct containing an N-terminal TEV-cleavable his-tag, which however also failed to crystallize.

Therefore, I used limited proteolysis to probe for flexible regions in the construct. Analytical digestion with trypsin at two different conditions (trypsin:Maf1 mass ratios of 1:500 or 1:5000) yielded distinct bands upon separation with SDS-PAGE (Figure 3.3). The digestion was repeated in a preparative scale and crystal trials set up with sample digested with a 1:500 and 1:5000 mass ratio. Needle-like crystals were obtained for the 1:500 sample, but not for the 1:5000 condition. However, these crystals did not diffract. To identify the domain boundaries that corresponded to the protease resistant fragments, SDS-PAGE bands were analyzed by tryptic digest mass spectrometry in the EMBL proteomics core facility. This revealed that in the 1:5000 sample, peptides corresponding to the acidic C-tail of Maf1 were found, while these were absent in the 1:500 sample which yielded crystals. Therefore, I cloned and expressed a truncation construct lacking the acidic C-terminal tail, Maf1 i.d.ΔC, which gave rise to needle-like crystals and microcrystalline material in several conditions of the Morpheus crystallization screen.

Crystallization conditions were optimized using several refinement screens in both hanging and sitting drops with the help of Brice Murciano (Figure 3.3). Crystal size could be improved, however all crystals exhibited a needle-like morphology. Crystals were tested at the ERSF Grenoble, but diffraction could only be observed upon high-dose exposure and was highly anisotropic, with spots visible along one axis approaching 3.7 Å, while diffraction along the perpendicular axis was limited to ~8 Å. Moreover, attempts to index the crystal lattice using either Mosflm [8] or XDS [96] were unsuccessful, most likely due to the presence of two or more lattices. This was apparent both by inspection of crystal morphology as well as the presence of a large number of "bad" diffraction spots when indexing was attempted. The results of the crystallization trials are summarized in Figure 3.3.

At this stage, the crystal optimization was not further pursued, as I had obtained a high-resolution EM map of the Maf1-Pol III complex in the meantime.

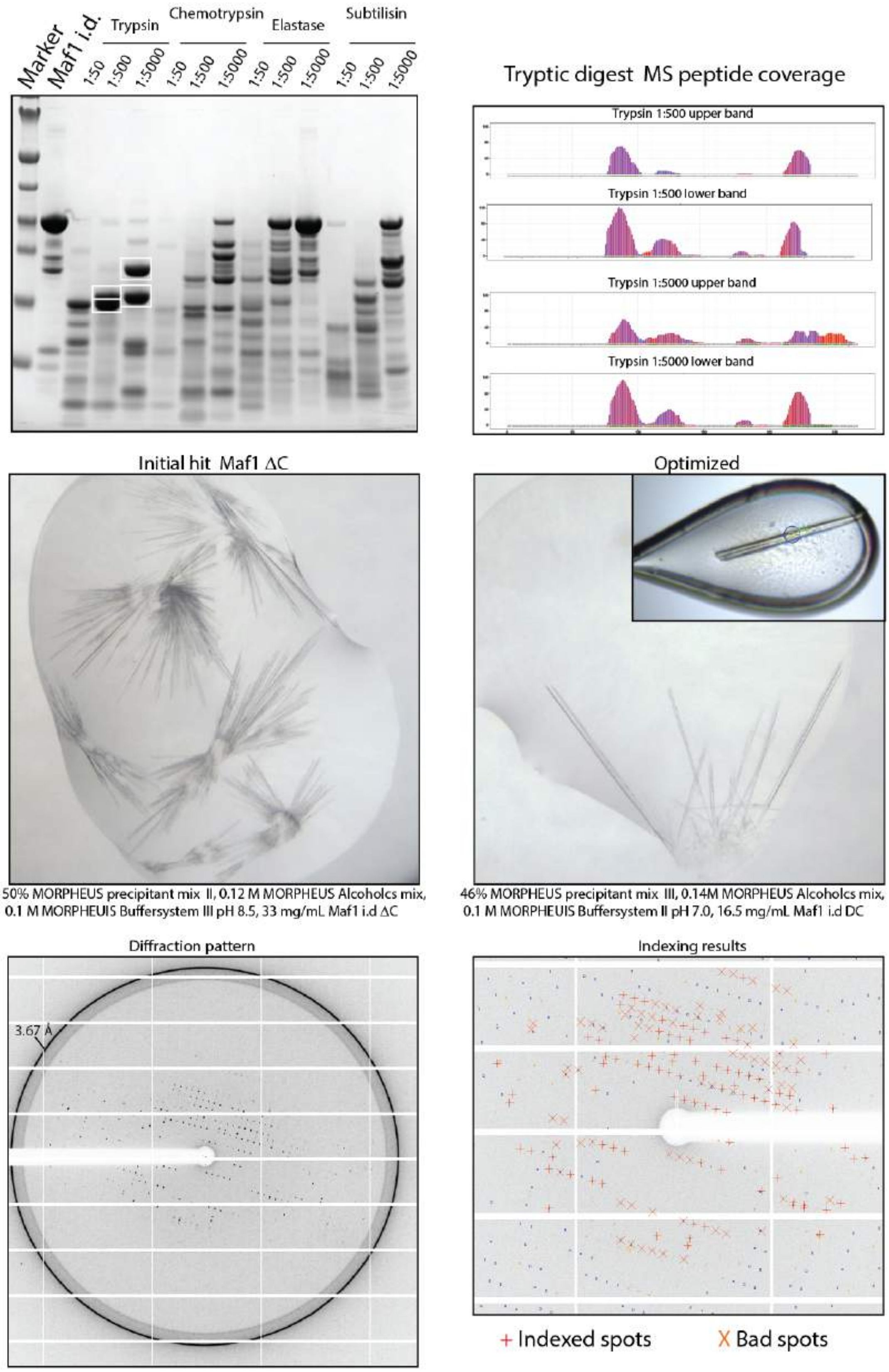

Marker
Maf1 i.d.
Trypsin
Chemotrypsin
Elastase
Subtilisin
1:50
1:500
1:5000
1:50
1:500
1:5000
1:50
1:500
1:5000
1:50
1:500
1:5000
Tryptic digest MS peptide coverage
Trypsin 1:500 upper band
Trypsin 1:500 lower band
Trypsin 1:5000 upper band
Trypsin 1:5000 lower band
Initial hit Maf1 ΔC
Optimized
50% MORPHEUS precipitant mix II, 0.12 M MORPHEUS Alcoholcs mix, 0.1 M MORPHEUIS Buffersystem III pH 8.5, 33 mg/mL Maf1 i.d ΔC
46% MORPHEUS precipitant mix III, 0.14M MORPHEUS Alcoholcs mix, 0.1 M MORPHEUIS Buffersystem II pH 7.0, 16.5 mg/mL Maf1 i.d DC
Diffraction pattern
3.67 Å
Indexing results
+ Indexed spots
X Bad spots

Figure 3.3: Crystallization of Maf1. From top to bottom: SDS-PAGE of limited proteolysis digests of Maf1 (left) and MS analysis of the bands indicated in white squares (right). Initial (left) and optimised (right) crystals of the Maf1ΔC construct. Diffraction pattern (left) and indexing results of a crystal with a 0.1 seconds exposure.

3.3 OPTIMIZATION OF THE MAF1-POL III SAMPLE FOR HIGH-RESOLUTION CRYO-EM

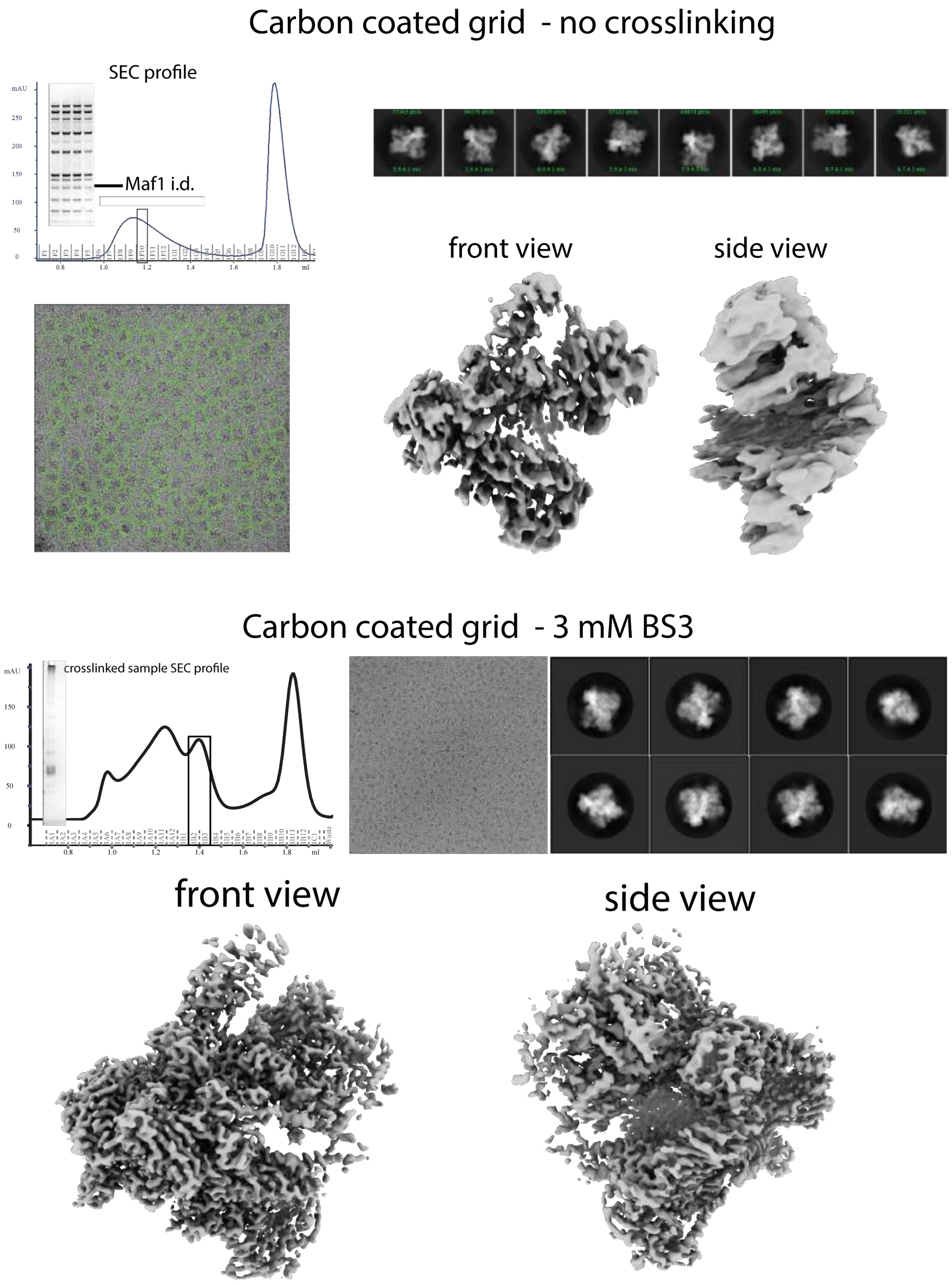
As mentioned before, the initial cryo-EM reconstruction showed very low occupancy of Maf1 and dissociation of the Pol III core from the heterotrimer and heterodimer. This is most likely caused by the cryo-EM sample preparation process, which involves blotting away excess liquid using filter papers, resulting in the formation of an ultra-thin liquid film. Both blotting and thin film formation tend to dissociate macromolecular complexes, because the shear forces present during blotting, and the extreme surface-to-volume ratio of a thin liquid film, where particles tend to adsorb to the air-water interface, often result in damaged particles or complex dissociation [64]. Moreover, evaporation rates are very fast for thin films, increasing the salt concentration in the sample, and mentioned as earlier, binding of Maf1 to Pol III is very sensitive to even moderate salt concentrations (> 50 mM Ammonium sulfate) *in vitro*.

To overcome these challenges, I initially tried to stabilize the complex by using grids covered with a thin carbon layer to which particles adsorb, reducing issues related to the air-water interface collision of particles. However, analysis of a large cryo-EM dataset did not show any density for Maf1 in the reconstructed maps, as well as a highly anisotropic resolution indicative of a strongly preferred orientation of the particles (Figure 3.4). I therefore decided to use chemical crosslinking, which should stabilize the complex and has the added benefit of altering the particles surface properties, often changing the orientation distribution. As my initial reconstruction showed Maf1 density relatively far away from Pol III, I decided to use (Sulfo-DSS)bis(sulfosuccinimidyl)suberate (BS3), as it has a longer crosslinking-distance then the previously used glutaraldehyde (11.4 Å [139] vs 7.5 Å [177]).

Crosslinking also allowed me to perform the binding reaction in low-salt conditions, but raise the salt concentration after crosslinking, which changes the elution profile of Pol III on size exclusion chromatography significantly. Pol III elutes in broad, early eluting peaks under low salt conditions, which might result from transient oligomerization. With higher salt concentrations, sharper peaks are obtained (Figure 3.4).

After crosslinking, the sample was fractionated using size exclusion chromatography and analyzed by negative stain EM, revealing three major populations: aggregated particles eluting in the void volume, a population of early eluting "clumping" particles, typically containg three to five Pol III complexes, and a late eluting monomeric fraction. I acquired negative stain datasets to probe the presence of Maf1 in the late eluting fraction, and the resulting reconstructions demonstrated that Maf1 was present (data not shown).

This sample preparation protocol was then used for initital cryo-EM reconstructions that revealed stabilization of Maf1 and isotropic resolution of the 3D reconstruction. Reconstructions of non-crosslinked and crosslinked samples are shown in Figure 3.4.



Carbon coated grid - no crosslinking

Carbon coated grid - 3 mM BS3

SEC profile
mAU
250
200
150
100
50
0
Maf1 i.d.
0.8
1.0
1.2
1.4
1.6
1.8
ml
front view
side view
crosslinked sample SEC profile
mAU
150
100
50
0
0.8
1.0
1.2
1.4
1.6
1.8
ml
front view
side view

Figure 3.4: Comparison of crosslinked and non crosslinked Maf1-Pol III reconstructions. Boxes in SEC profiles indicate the fractions used for grid preparation.

After optimizing blotting parameters to obtain thin ice, a large (10 524 micrographs) dataset was acquired on a Titan Krios equipped with a K2 detector operated in counting mode, which allowed the reconstruction of a 3.25 Å map from 117 442 particles. In this map, the Pol III core is resolved at better than 3 Å, and the local resolution for Maf1 density ranges from 3.6 Å to 4.3 Å, which allowed building of an atomic model (Figure 3.7).

3.4 PARTICLE CLASSIFICATION OF THE MAF1 DATASET

Due to the large number of particles in the dataset (1 6788 795), the dataset was divided into four batches and each was classified by 3D classification in RELION 3.0 using a 60 Å low-pass filtered model of apo Pol III (PDB 5fj9) as reference. The best class of each batch was retained, and all batches were combined, yielding a dataset of 728 000 particles, and refined. The resulting map showed density for Maf1 and an adjacent density corresponding to C34 WH2, but at a lower threshold. Using a classification focused on this region, Maf1 and C34-WH2 densities could be improved (305 000 particles). Finally, a masked classification focussing on the stalk, trimer and clamp seperated an open clamp state, and a closed clamp state, containing 117 442 particles. Density for Maf1 was smeared out in the open clamp state, but clear separation of β-sheets and sidechain density were visible in the closed clamp state (Figure 3.5).

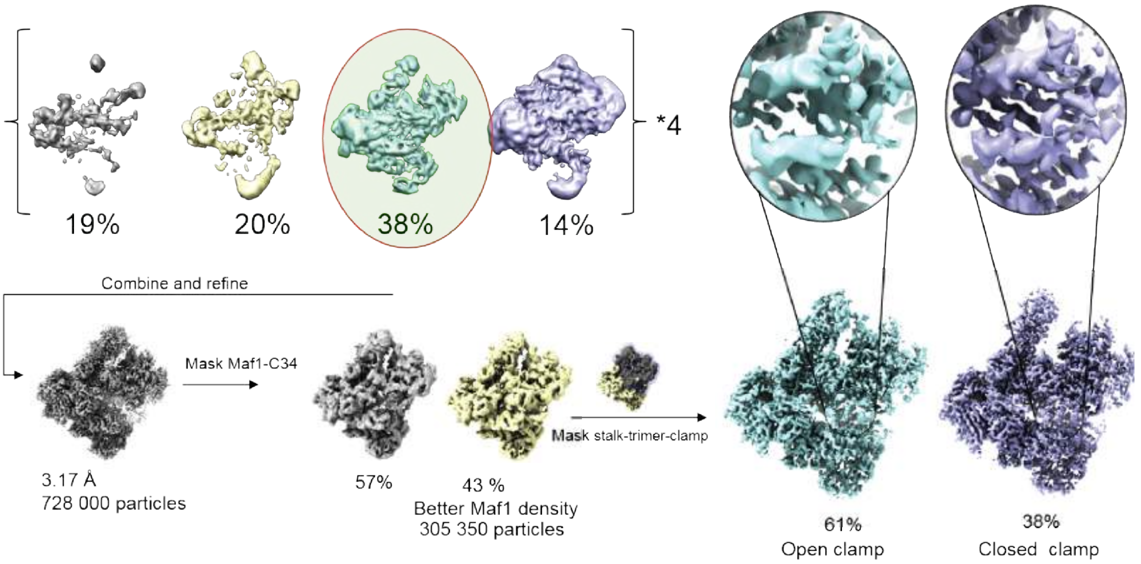


Figure 3.5: Particle classification of the Maf1-Pol III dataset. The dataset was divided into four batches which were classified in 3D and the best class of each batch was retained. Batches were combined, refined and sequentially classified using masks on Maf1 and the stalk-heterotrimer-clamp region. Blow-ups show density for Maf1 in the open clamp and closed clamp states.

Starting from the 305 000 particle set, I also performed MultiBody refinement implemented in RELION 3, using two masks. The first covered the core of Pol III and the heterodimer, and the second covered the stalk, clamp, heterotrimer and Maf1. This significantly improved side-chain density in the heterotrimer and clamp (Figure 3.6). Therefore, I used the maps obtained in this way to improve the model of Pol III in this region (starting from PDB 6eu3) and corrected minor sequence register shifts. Progress of model building was monitored by real-space refinement of partial models, namely the heterodimer, the heterotrimer-clamp region, and Maf1 against respective maps obtained from MultiBody refinement.

Heterodimer	**Before**	**after**
MolProb score	2.16	1.89
Clash score	7.22	5.29
Rotamer outliers %	1.58	0
Ramachandran outliers %	0.35	0

Trimer-clamp:	**Before**	**after**
MolProb score	2.9	1.84
Clash score	14.36	5.14
Rotamer outliers %	4.76	0.39
Ramachandran outliers %	1.08	0.07

Maf1	**after**
MolProb score	1.89
Clash score	5.29
Rotamer outliers %	0.0
Ramachandran outliers %	0.0

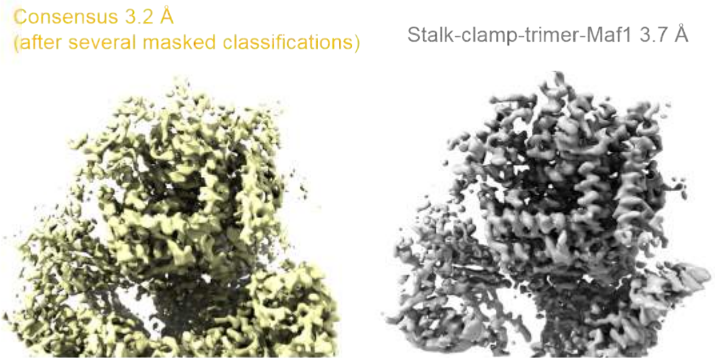

Consensus 3.2 Å
(after several masked classifications)
Stalk-clamp-trimer-Maf1 3.7 Å

Figure 3.6: Multibody refinement of Maf1-Pol III. Densities from the consensus refinement (yellow) and Multibody refinement (grey) of the heterotrimer-clamp regions are compared. Tables show refinement and model statistics of the indicated regions before and after adjustment of the model coordinates using the Multibody-refinement maps.

3.5 STRUCTURE OF THE MAF1-POL III COMPLEX

The structure of yeast Maf1 is very similar to that of human Maf1 [201] with a central β-sheet sandwiched by three helices on one side (interface A) and a single helix on the other side (interface B). Maf1 binds Pol III between the protrusion domain, the wall domain and the clamp coiled-coil helices, with its interface A oriented towards the wall (Figure 3.7). A prominent tryptophan residue (W294 in C160) of the clamp coiled-coil inserts into a central furrow of Maf1, where it stacks with Maf1 W319. Interestingly, C160 W294 is part of a Pol III specific highly conserved 'template strand pocket' [1] involved in DNA unwinding and transcription initiation. Maf1 W319 is also invariantly conserved (Figure 3.8), suggesting that the aromatic stacking interaction is an important determinant of Maf1 binding.

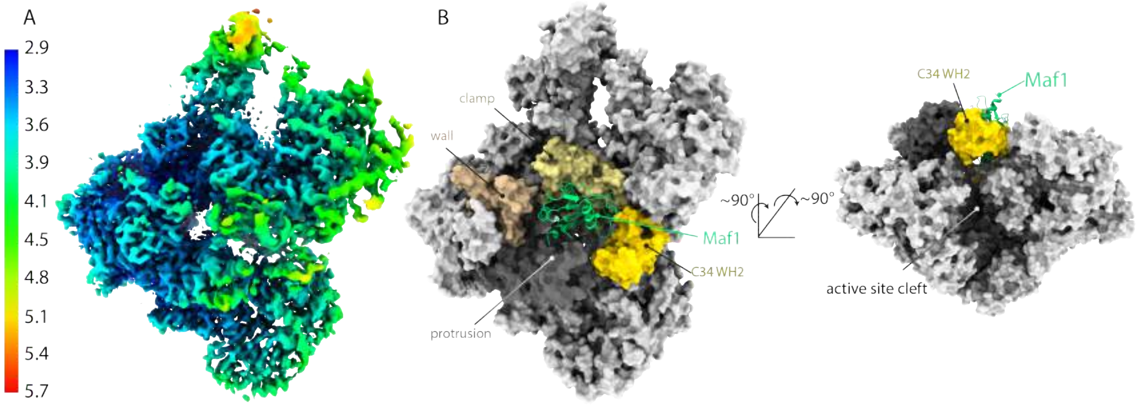
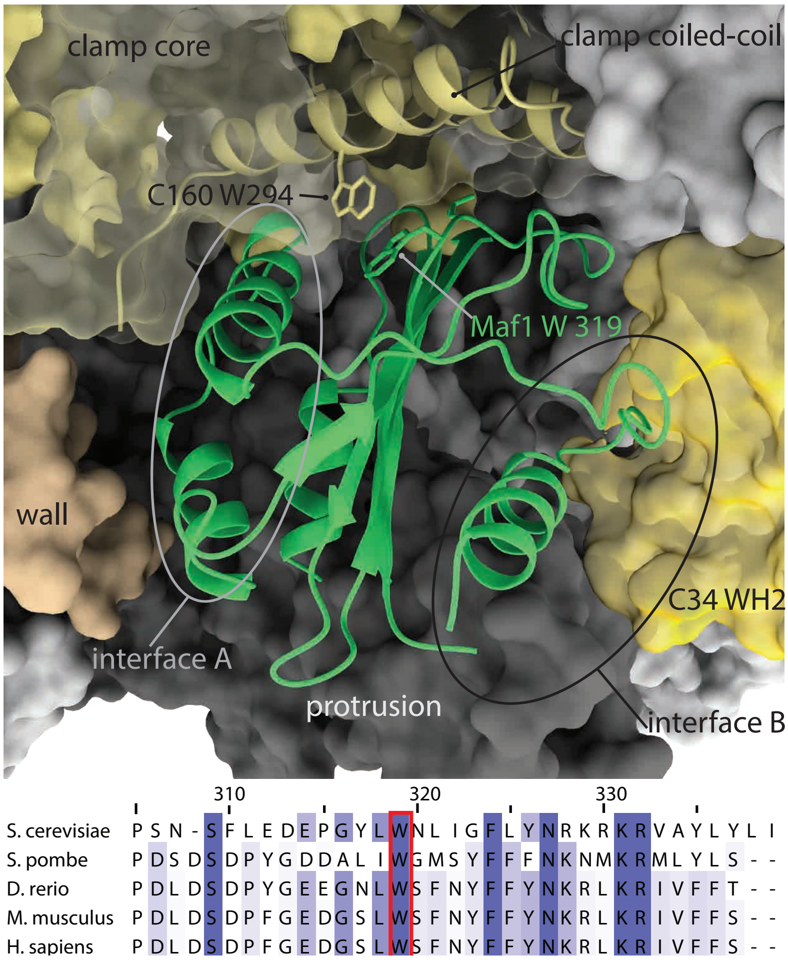


Figure 3.7: Structure of the Maf1-Pol III complex. A: Cryo-EM map of the Maf1-Pol III complex colored by local resolution. B: Structure of the Maf1-Pol III complex, with Pol III shown as surface and Maf1 shown as ribbons in two different views. Interacting domains in Pol III are labelled.

On the opposite site of the wall, the mobile C34 WH2 domain becomes ordered and interacts with Maf1 interface B. To confirm this interaction, I performed pulldown experiments of purified Maf1 and his-tagged C34 (1-156), which independently demonstrates binding of Maf1 to this construct (Figure 3.10).

Mapping the electrostatic potential onto the surface of Maf1 reveals a striking feature of Maf1 interface B. Two loops, namely residues 245-256 and residues 307-319, are highly positively charged. Superposition with the structure of the Pol III pre-initiation complex in the open state (PDB 6f40) further shows that the backbone of promoter DNA partially overlaps with these loops, and that the interface B is a near-perfect mimick of one turn of B-DNA (Figure 3.10). The C34 WH2 binds this interface by inserting its positively charged 'wing' between the two loops.

Concomitant with the ordering of the C34 WH2, Pol III adopts a closed clamp state, resembling the pre-initiation complex (Figure 3.9).


clamp core
clamp coiled-coil
C160 W294
Maf1 W 319
wall
C34 WH2
interface A
protrusion
interface B
310
320
330

S. cerevisiae	P S N - S F L E D E P G Y L W N L I G F L Y N R K R K R V A Y L Y L I
S. pombe	P D S D S D P Y G D D A L I W G M S Y F F F N K N M K R M L Y L S - -
D. rerio	P D L D S D P Y G E E G N L W S F N Y F F Y N K R L K R I V F F T - -
M. musculus	P D L D S D P F G E D G S L W S F N Y F F Y N K R L K R I V F F S - -
H. sapiens	P D L D S D P F G E D G S L W S F N Y F F Y N K R L K R I V F F S - -

Figure 3.8: Close-up of the Maf1-Pol III interaction. Two stacking Trp-residues are highlighted in stick representation. Bottom: Sequence alignment of Maf1, showing that the stacking Trp residue 319 is invariantly conserved.

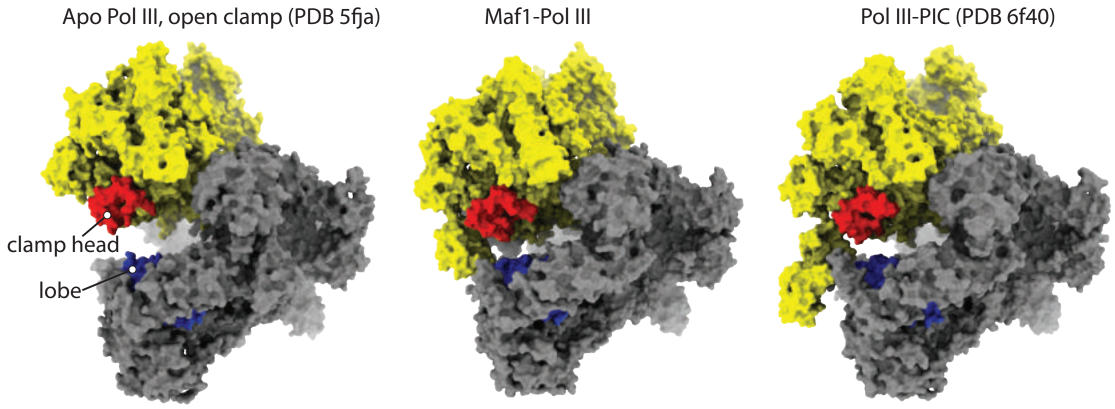

Apo Pol III, open clamp (PDB 5fja)
Maf1-Pol III
Pol III-PIC (PDB 6f40)
clamp head
lobe

Figure 3.9: Side-by-side comparison of the Pol III conformation in the open clamp state, the Maf1 bound state and the pre-initiation state. The stalk-heterotrimer-clamp region is coloured yellow, the clamp head is colored red, and the lobe is coloured blue. The distance between the clamp head and lobe indicate if Pol III is in an 'open clamp' or a 'closed clamp' state.

3.6 mechanism of repression by maf1

The binding site of Maf1 overlaps with the location of TFIIIB and promoter DNA in the PIC (Figure 3.11). This explains the finding that binding of Maf1 and TFIIIB are mutually exclusive [201] and provides a mechanism for transcription repression by Maf1. While previous work [201] has positioned Maf1 on top of the clamp coiled-coil and explained the mechanism of Pol III repression with allosteric repositioning of the C82/C34/C31 heterotrimer subcomplex, the structure presented here suggests a direct competition for the same binding interface. Moreover, Maf1 induces allosteric changes that closely mimic the PIC state, and sequesters residues that are involved in establishing contacts with promoter DNA in the PIC. Specifically, the protrusion, the template strand pocket and the C34 WH2 "wing" are all bound by Maf1, and the active site cleft is sealed off due to ordering of the C34 WH2 (Figure 3.7).

Although the acidic C-terminus of Maf1 is disordered, the way Maf1 is orientated suggests that the C-terminus projects into the active site cleft. This might help to repel nucleic acids electrostatically and thereby prevent recruitment to open chromatin. Such a mechanism would also explain why the sequence of the *S. cerevisiae* C-terminus can be swapped for that of *S. pombe* or human Maf1 without affecting function (I. Willis, personal communication). The sequence of this C-terminal does not show any strict conservation but is in all cases enriched in negatively charged residues.

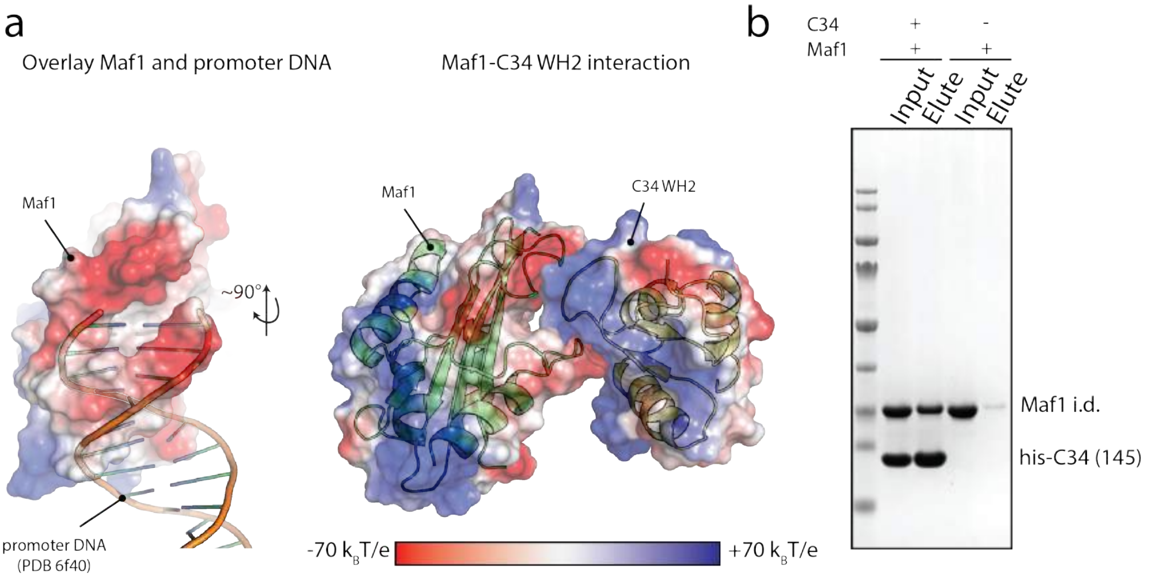


Figure 3.10: Maf1 mimics promoter DNA to bind C34. a: Overlay of Maf1 coloured by electrostatic potential with promoter DNA from PDB 6f40 revealed that Maf1 mimicks B-DNA in shape and charge. The 'wing' of C34 WH2 inserts between the 'backbone-mimick loops' of Maf1. b: recombinant Maf1 binds recombinant C34(1-156) in *in vitro* pull-down experiments.

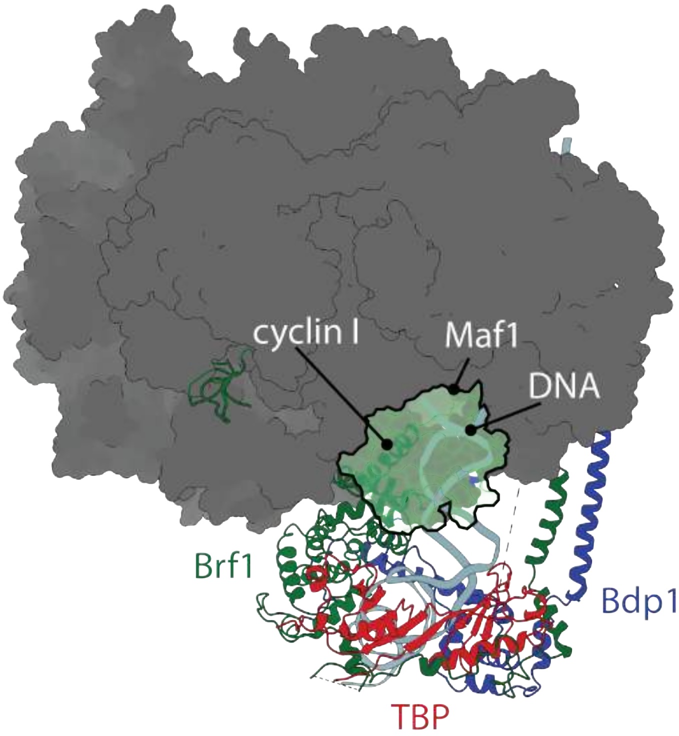


Figure 3.11: Maf1 blocks the TFIIIB binding site. Superposition of the Maf1-Pol III structure with the Pol III pre-inititation complex (PDB 6f40). Maf1 is shown as a transparent surface and outlined with a thick black line, TFIIIB and promoter DNA are displayed in ribbon representation.

.

4

ASSEMBLY OF TFIIIB BY TFIIIC

The structure of TFIIIC is a long-standing interest of the Müller lab, and cryo-EM characterization of TFIIIC was initiated by former postdoc Heena Khatter (H.K.). I joined the project in the third year of my PhD, and have continued the work since H.K. has left the laboratory in August 2018 with the help of Anna Jungblut (A.J.), Master student in the lab since December 2018.

4.1 RECOMBINANT EXPRESSION OF TFIIIC

Endogenous TFIIIC has been purified in our lab from large amounts of yeast cells obtained by fermentation. This method yielded pure and homogeneous protein, but is costly and yields are very low (100-250 µg from 150 L), which presented a major bottleneck in the structural characterization of TFIIIC.
Therefore, former lab members used two approaches for recombinant expression in insect cells:

- Expression from a construct in which all genes are cloned into the same open reading frame, where individual genes are separated by TEV cleavage sites to ensure stoichiometric expression of all six subunits
- Generation of two viruses encoding for the τA subcomplex and τB subcomplexes, respectively. Each gene is controlled under its own promoter (either the polyhedron (polH) or the p10 promoter) and terminator elements, and insect cells are co-infected with two viruses.

Neither of these strategies proved successful. Therefore, we set out to express all six subunits from the same virus, with each gene being under the control of the polH promoter. This project was initiated together with H.K. and cloning and virus generation were performed with the help of Helga Groetsch.
While expression of all six subunits using this strategy was confirmed early on using tryptic digest mass spectrometry of SDS-PAGE bands and western blotting, the purification protocol had to be optimized for yield and homogeneity, as certain subcomplexes of TFIIIC were over-represented after affinity purification over Ni^{2+}-NTA and Strep-tactin resins (Figure 4.1).
In particular, cleavage of the tags with TEV protease immediately after affinity purification resulted in the formation of a precipitate that quantitatively removed all TFIIIC and the TEV protease from solution. When TEV cleavage was performed after additional purification steps that removed the excess of the τ60 and τ91 subunits, this precipitation was still observed but to a lesser extent. It is therefore likely that the excess subunits become destabilized upon TEV cleavage and trigger co-precipitation of other proteins in the solution.
To test if TEV cleavage was required to obtain functional TFIIIC, we performed an *in vitro* transcripton assay comparing endogenous TFIIIC purified from yeast with recombinant

TFIIIC before and after TEV cleavage. This showed that tagged, recombinant TFIIIC had a similiar level of activity as endogenous TFIIIC, but TEV cleaved recombinant TFIIIC had a lower level of activity (Figure 4.2). Therefore, the TEV cleavage step was omitted from subsequent purifications.

In terms of yield, a certain variability was observed. The best purification yielded 2 mg of TFIIIC from 6L of high five insect cell expression culture, however 80% of that precipitated upon TEV cleavage. Other batches yielded about 0.5 mg tagged TFIIIC from 3L. This represents a significant improvement over the endogenous purification strategy in terms of yields, cost and labor.

To further optimize the expression system, different combinations of tags are currently being tested as part of the Master thesis project of (A.J.), with the aim to simplify the purification, ideally yielding stoichiometric holo-TFIIIC after the affinity purification step. An important advantage of the insect cell expression system is that it enables us to express subcomplexes of TFIIIC separately, as shown in section 4.3.

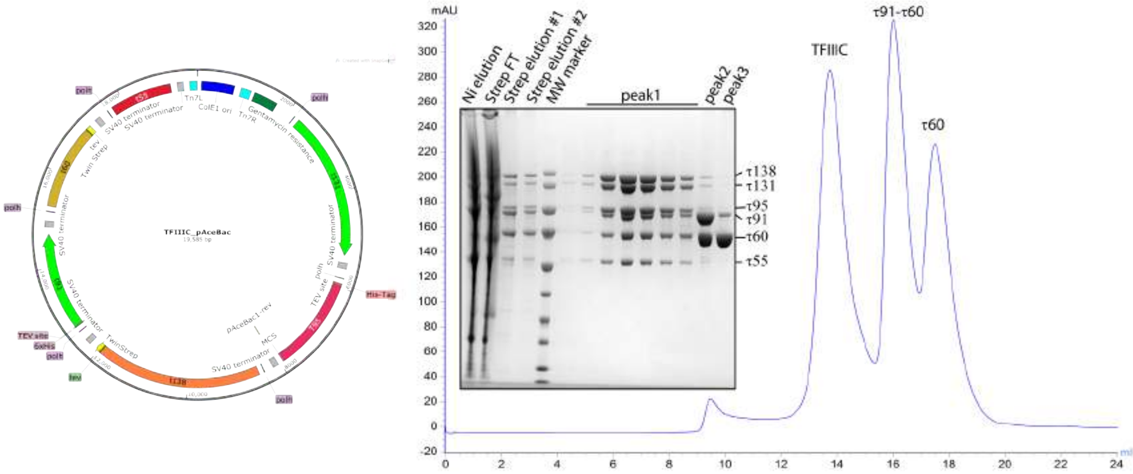


Figure 4.1: Insect cell expression of TFIIIC. The expression vector used for bacmid generation is shown on the left. SDS-PAGE of the purification and a chromatogram of a Superose 6 elution of TFIIIC are shown on the right. τ60-Strep and his-τ91 are expressed in excess of the other subunits and have to be removed after affinity purification.

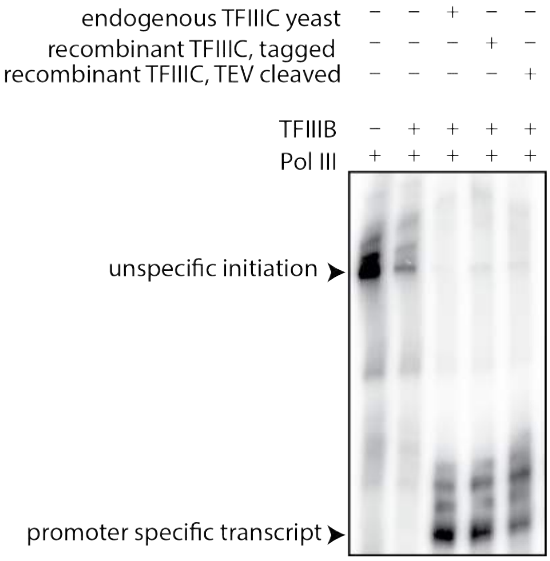

endogenous TFIIIC yeast − − + − −
recombinant TFIIIC, tagged − − − + −
recombinant TFIIIC, TEV cleaved − − − − +
TFIIIB − + + + +
Pol III + + + + +
unspecific initiation
promoter specific transcript

Figure 4.2: Recombinant TFIIIC is active. Comparison of the promoter specific *in vitro* transcription levels with endogenous or recombinant TFIIIC. The product labelled "unspecific initiation" likely results from initiation at free 3′ overhangs. Cropped image. Experiment performed by F.B.

4.2 EFFORTS TO RECONSTITUTE THE TFIIIC-TFIIIB-POL III COMPLEX

Previous work in the lab by H.K. had focused on obtaining cryo-EM maps of free or DNA-bound TFIIIC. However, this had not resulted in interpretable reconstructions, presumably due to the intrinsically flexible nature of TFIIIC.

Therefore, we sought to rigidify the complex by locking it in a defined conformation by binding it to interaction partners. Together with H.K., significant effort was invested into reconstituting the TFIIIC-TFIIIB-Pol III complex. While co-elution of TFIIIC, TFIIIB and Pol III could be detected from SEC experiments, interpretation of this was difficult due to the similar molecular weight of Pol III and TFIIIC. Moreover, Pol III tends to aggregate and form dimeric and higher oligomeric state assemblies in the presence of low salt and nucleic acids, as evident by negative stain EM. Therefore, particles appearing large by visual inspection on negative stain grids—as expected for a complex between Pol III and TFIIIC—cannot unambiguously be classified as being TFIIIC-bound. In summary, five different DNA scaffolds and three different crosslinking agents (glutaraldehyde, BS3 and formaldehyde) as well as various buffer compositions were tested in hope of promoting complex formation. However, negative stain analysis of several datasets consistently revealed the presence of unbound Pol III, Pol III dimers, the Pol III-TFIIIB complex and 2D classes corresponding to free TFIIIC, but no classes consistent with a Pol III-TFIIIB-TFIIIC complex were detected (exemplified in Figure 4.3).

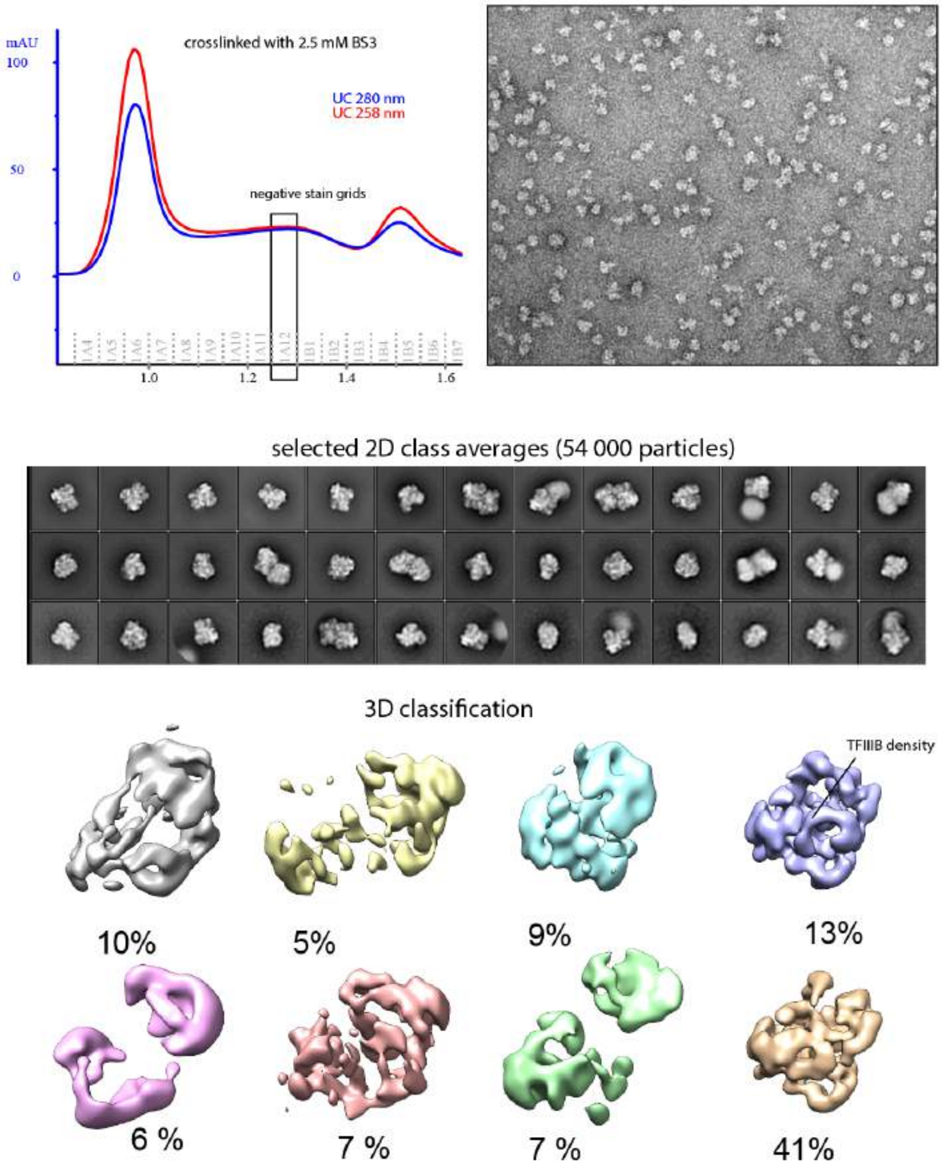


Figure 4.3: Negative stain data analysis of a sample of TFIIIC, TFIIIB and Pol III, crosslinked with 2.5 mM BS3. Top: chromatogram of a Superose 6 INCREASE 3.2/300 column fractionation and representative negative stain micrograph of the indicated fraction. Middle: Selected 2D classes, revealing the presence of Pol III dimers. Bottom: 3D classification using Pol III as a reference, indicating that unbound Pol III and the Pol III-TFIIIB complex are the major populations.

4.2.1 *Efforts to reconstitute the TFIIIC-TFIIIB-TFIIIA-Pol III complex on a type I promoter*

Together with A.J., we also set out to reconstitute the type I promoter system to test if it is better behaved for structural studies. This was motivated because studies from the Pol II field report increased stability of reconstituted transcription complexes with increasing amount of transcription factors (personal communication).

For this end, TFIIIA was cloned and successfully purified from *E. coli*, and the long (178 bp) DNA fragment containing the 5S rDNA gene and upstream promoter sequence was produced via large scale PCR and purified by ion exchange chromatography. However, binding analysis suggested that TFIIIA becomes displaced from the DNA when TFIIIC is recruited, suggesting that it might serve as an assembly factor rather than as an integral part of the PIC (data not shown).
We therefore shifted our focus to the characterization of a complex between TFIIIC and TFIIIB in absence of Pol III.

4.2.2 *Reconstitution of a TFIIIB-TFIIIC complex*

To obtain a well-behaved, stable sample, I first sought to identify a suitable promoter sequence for reconstitution by testing candidate sequences in an *in vitro* transcription assay with purified transcription factors and Pol III. I selected five candidate sequences, based on the following criteria: two sequences (His-tRNA and Met-tRNA) were reported to be highly enriched in TFIIIC chromatin immunoprecipitation (ChIP) experiments [75, 146]; one tRNA sequence was reported to be transcribed independently of TFIIIC *in vitro* due to a strong TATA box [42]; a synthetic construct was designed which combines the upstream promoter elements of the U6 promoter with a downstream Glu-tRNA gene sequence, and the 5s rDNA gene requiring TFIIIA, to test another promoter type. Of these, the His-tRNA sequence was notably stronger transcribed then all others, and was chosen for all subsequent experiments (Figure 4.4).

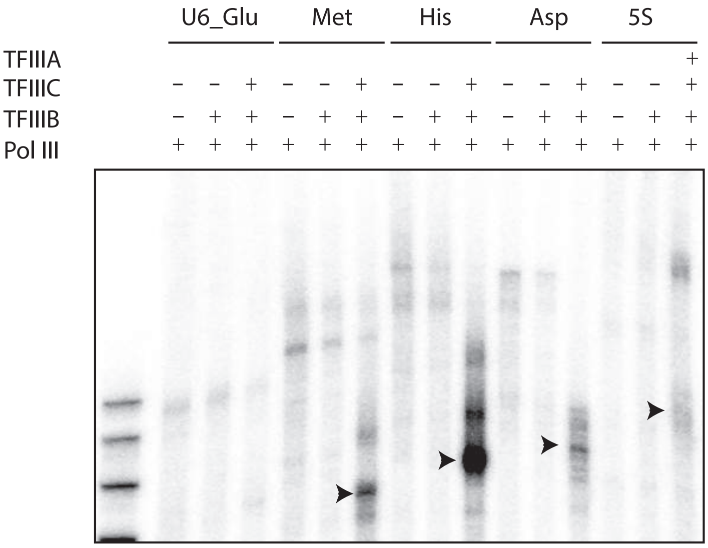


Figure 4.4: *In vitro* transcription of various Pol III promoter sequences. Specific transcripts are indicated by arrowheads, whereas longer transcripts presumably origin from transcription initiated on free 3′ ends. Cropped image. Experiment performed by F.B.

4.3 STRUCTURAL ANALYSIS OF THE τA SUBCOMPLEX

As already mentioned, attempts to determine the structure of holo-TFIIIC had proven fruitless. Reasoning that the flexible nature of TFIIIC, with the two modules τA and τB that can accommodate promoter elements separated by anything from 31 to 93 base pairs, might impair accurate alignment of particles and therefor structure determination. I decided to express the τA and τB modules of TFIIIC separately in insect cells. The τA module is the most conserved part of TFIIIC [86], and also responsible for interaction with TFIIIB and Pol III, and we hence focused our attention on τA first.

Viruses carrying expression cassettes for the τA subunits were generated by A.J. τA could be efficiently expressed and purified from insect cells, yielding about 3 mg of protein from 3L of culture. The protein was tested for functionality in two assays: First, an *in vitro* transcription assay showed that τA can support transcription from a TFIIIC-dependent promoter (the histidine tRNA gene), although at lower levels then full-length TFIIIC (Figure 4.5). Second, τA efficiently bound TFIIIB when incubated together with a DNA oligo encompassing the A-box and 35 bp upstream of the transcription start site (Figure 4.5).

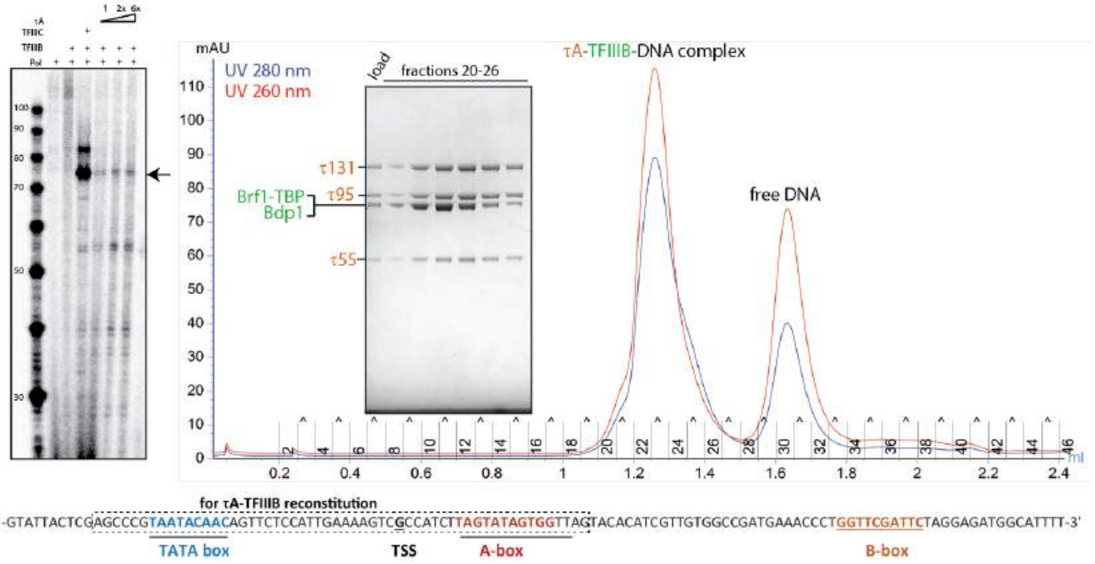
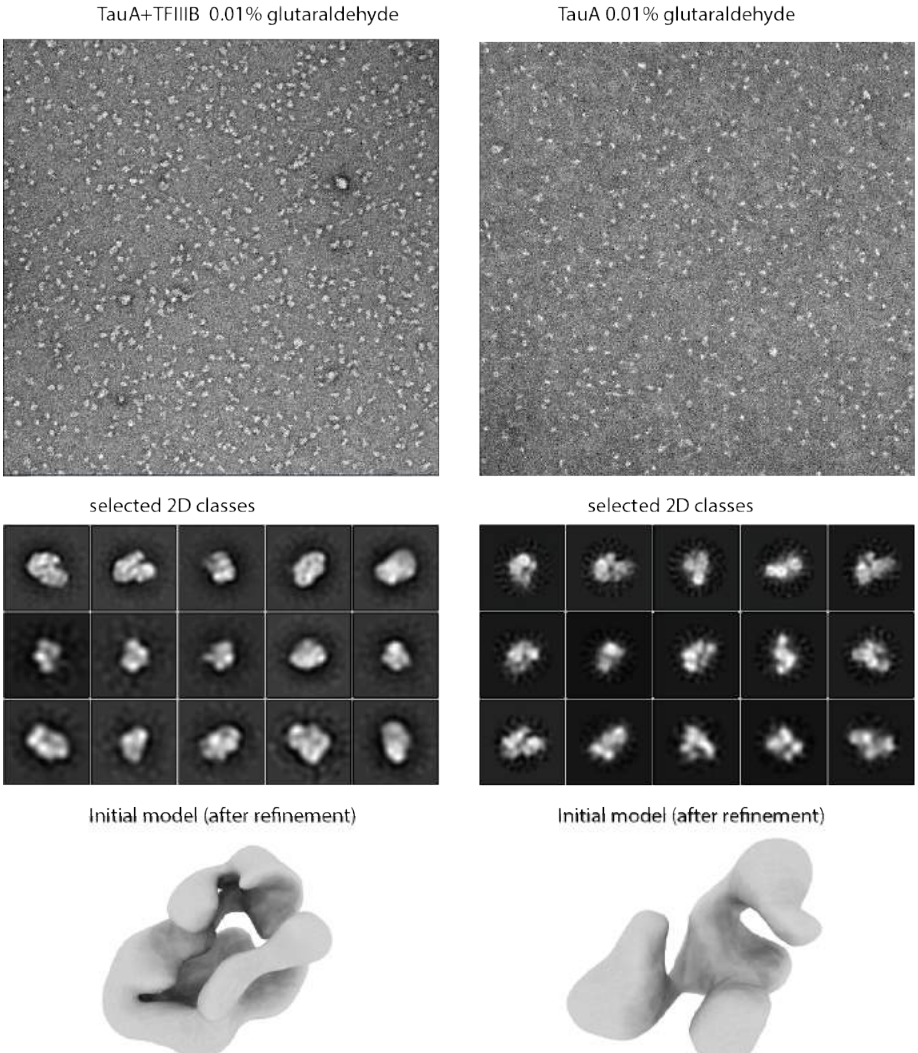


Figure 4.5: τA stimulates basic levels of transcription and stoichiometrically binds TFIIIB. *In vitro* transcription of the His_tH(GUG)E2 gene using either endogeneous TFIIIC or recombinant τA. Right: SEC profile of τA and TFIIIB on a Superdex 200 column. Bottom: sequence of the non-template strand of the HistH(GUG)E2 DNA oligo used for transcription assays. The dotted box delineates the sequence used for the reconstitution experiment.

τA and the τA-TFIIIB complex were further analysed by negative stain EM. While τA appeared heterogeneous and yielded noisy 2D classes, the τA TFIIIB dataset gave rise to featured 2D classes, which were selected to obtain a homogeneous particles set as input for the RELION initial model program (Figure 4.6). The resulting model was used as a reference to classify the entire dataset (see chapter 6 for details), and a reconstruction with interpretable density was obtained. This allowed unambiguous fitting of TFIIIB, as well as the crystal structure of the τ131 N-terminal TPR array (Figure 4.8). Moreover, additional densities that fitted in dimension with two other available crystal structures, namely a dimeric τ95-τ55 fragment and a τ95 DNA binding domain were present (Figure 4.8). An additional density that could potentially be occupied by the C-terminal TPR array of τ131 is also present adjacent to the C-terminus of the N-terminal TPR (Figure

4.8). The negative stain map shows a central channel that connects from TBP to the back of the molecule, which is a potential path of promoter DNA (which generally does not show density in negative stain reconstructions). Moreover, the architecture of the complex derived from these density fits is in agreement with the topology of the TFIIIB-τA complex, where TFIIIB binds upstream of the TSS and τA occupies the downstream A-box. In addition, the N-terminal τA-TPR array is close to both TFIIIB subunits, in agreement with biochemical evidence presented in section 1.3.3.3, giving confidence in the validity of this reconstruction and density fits. However, these fits will have to be validated by high-resolution EM maps.



TauA+TFIIIB 0.01% glutaraldehyde

TauA 0.01% glutaraldehyde

selected 2D classes

selected 2D classes

Initial model (after refinement)

Initial model (after refinement)

Figure 4.6: Negative stain analysis of the τA-TFIIIB complex and τA only. While the TFIIIB-τA complex gives featured 2D classes, apo τA appears highly flexible and yields an uninterpretable 3D reconstruction

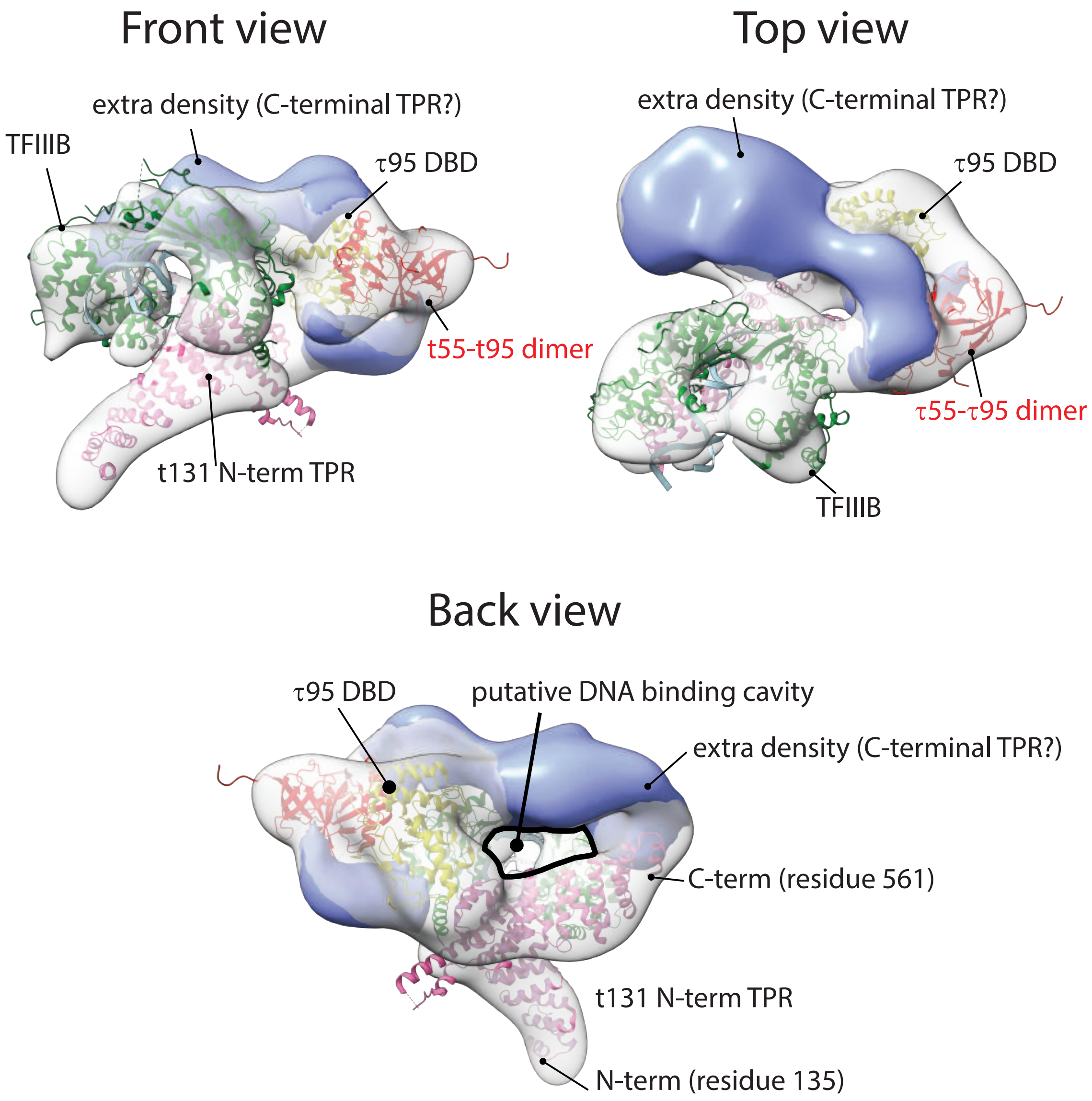

Front view
extra density (C-terminal TPR?)
TFIIIB
τ95 DBD
t55-t95 dimer
t131 N-term TPR
Top view
extra density (C-terminal TPR?)
τ95 DBD
τ55-τ95 dimer
TFIIIB
Back view
τ95 DBD
putative DNA binding cavity
extra density (C-terminal TPR?)
C-term (residue 561)
t131 N-term TPR
N-term (residue 135)

Figure 4.7: Negative stain density fit of τA and TFIIIB. Fits for τ95-τ55 (PDB 4bji) and τ95 DBD (PDB 4bjj) are not unambiguous from the density but were placed based on additional considerations: the τ95 DBD was positioned such that the predicted DNA binding interface faces the central channel, and the τ95-τ55 dimer was positioned close to the DBD and fitted using the UCSF chimera fit in map program. The density not explained by the fitted (crystal) structures is coloured in blue, and might correspond to the predicted C-terminal TPR array of τ131.

4.3.1 *Towards a high resolution cryo-EM structure of the τA-TFIIIB complex*

Motivated by having obtained the first interpretable map of a TFIIIC module, we further pushed towards obtaining a high resolution cryo-EM structure.
An initial dataset was collected from a sample that was prepared by mixing components in 200 mM potassium acetate (KOAc) and crosslinking with glutaraldehyde, following the protocol used for negative stain sample preparation. Cryo-EM grids were prepared with Cu 2/1 grids covered by a 2 nm Carbon layer. A small dataset was collected with a high

dose to improve low-resolution particle contrast. However, 2D classification of particles did not reveal distinct shapes that would resemble those seen in negative stain classes. This might be caused by extensive inter-particle crosslinks, which is favoured by the ten times higher protein concentration at which the cryo sample was crosslinked.

In parallel, sample homogeneity could be improved by lowering the ionic strength to that of 50 mM KOAc, which led to improved binding of the DNA scaffold as observed on SEC separations. A second cryo-EM dataset of an uncrosslinked sample that was purified via SEC in 50 mM KOAc was collected on the same grid type. While this yielded some 2D classes that showed distinct shapes, they were of low resolution and did not resemble 2D classes from the original negative stain dataset of crosslinked sample. A second negative stain dataset of the same sample that was used for the second cryo dataset showed similar 2D classes as observed in cryo, indicating that the τA-TFIIIB complex requires crosslinking to be amenable for EM preparation.

A third dataset, collected from gently crosslinked sample in 50 mM KOAc and subsequent purification by SEC however also did not result in interpretable 2D classes.

Apart from sample homogeneity, a second factor that might impair image alignment is the poor contrast resulting from the carbon layer. Therefore, grids without carbon support were also tested. The sample looked highly aggregated, presumably due to denaturation at the air-water interface.

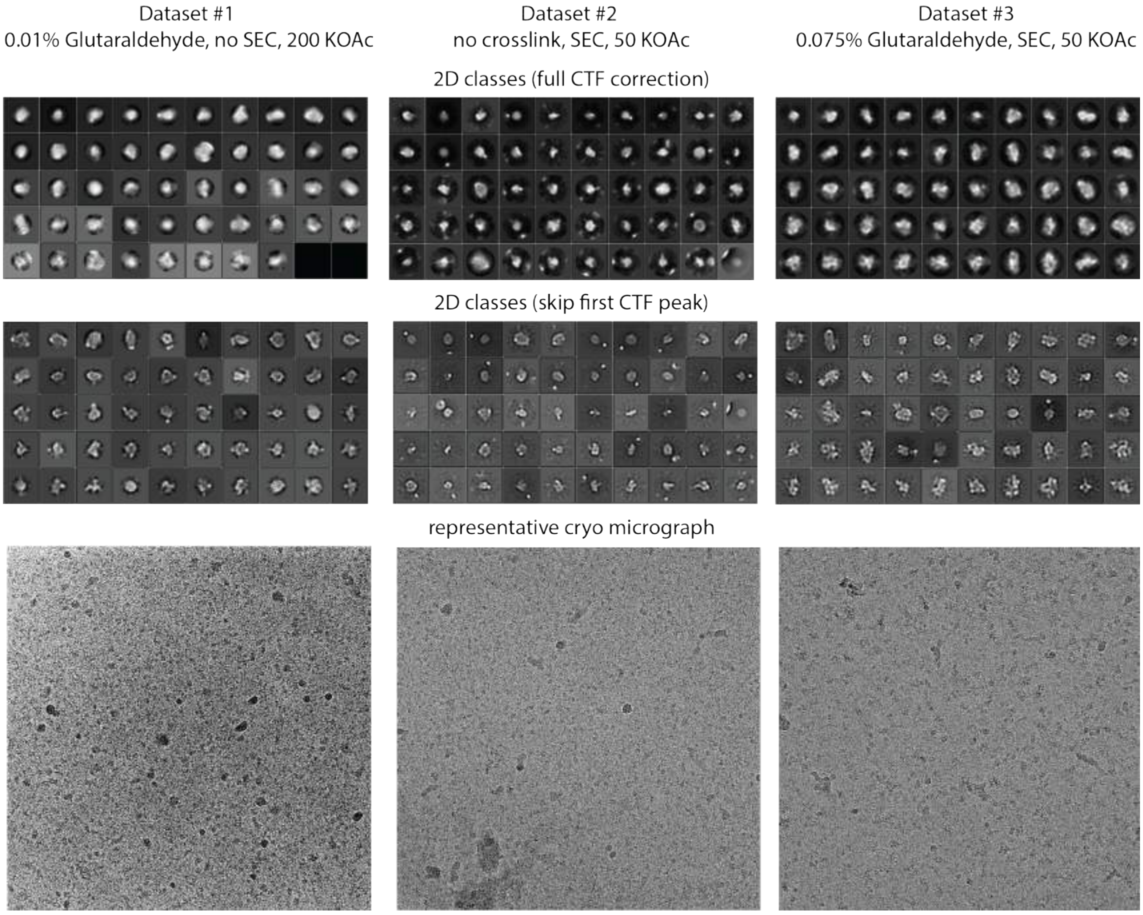

Dataset #1
0.01% Glutaraldehyde, no SEC, 200 KOAc
Dataset #2
no crosslink, SEC, 50 KOAc
Dataset #3
0.075% Glutaraldehyde, SEC, 50 KOAc
2D classes (full CTF correction)
2D classes (skip first CTF peak)
representative cryo micrograph

Figure 4.8: 2D classes of cryo datasets collected of τA and TFIIIB. The 50 most highly populated classes of 2D classifications runs in RELION with either fully corrected CTF (top) or with the RELION option 'skip first CTF peak' are shown for the datasets indicated above.

4.3.2 *Comparison of the τA-TFIIIB map with negative stain maps of TFIIIC and TFIIIC-TFIIIB*

In order to assign the τA module in the TFIIIC holo-complex, I reprocessed negative stain data of TFIIIC bound to DNA and a TFIIIC-TFIIIB-DNA complex[1]. While featured 2D classes were obtained, it was not possible to unambiguously fit the τA density or the TFIIIB density in any of the classes. However, it was striking that the TFIIIB-TFIIIC complex appeared much more elongated in both 2D and 3D classes compared to the TFIIIC complex (largest dimension 260 Å vs 180 Å, see Figure 4.9).

Given that Bdp1 competes for a binding spot on τ131 as the τIR of τ138, this elongated shape might be the result of a large scale conformational change in TFIIIC, triggered by release of the τIR-131 interaction. This might help to render the complex initiation competent.

---

[1] Sample preparation and data collection by H.K.

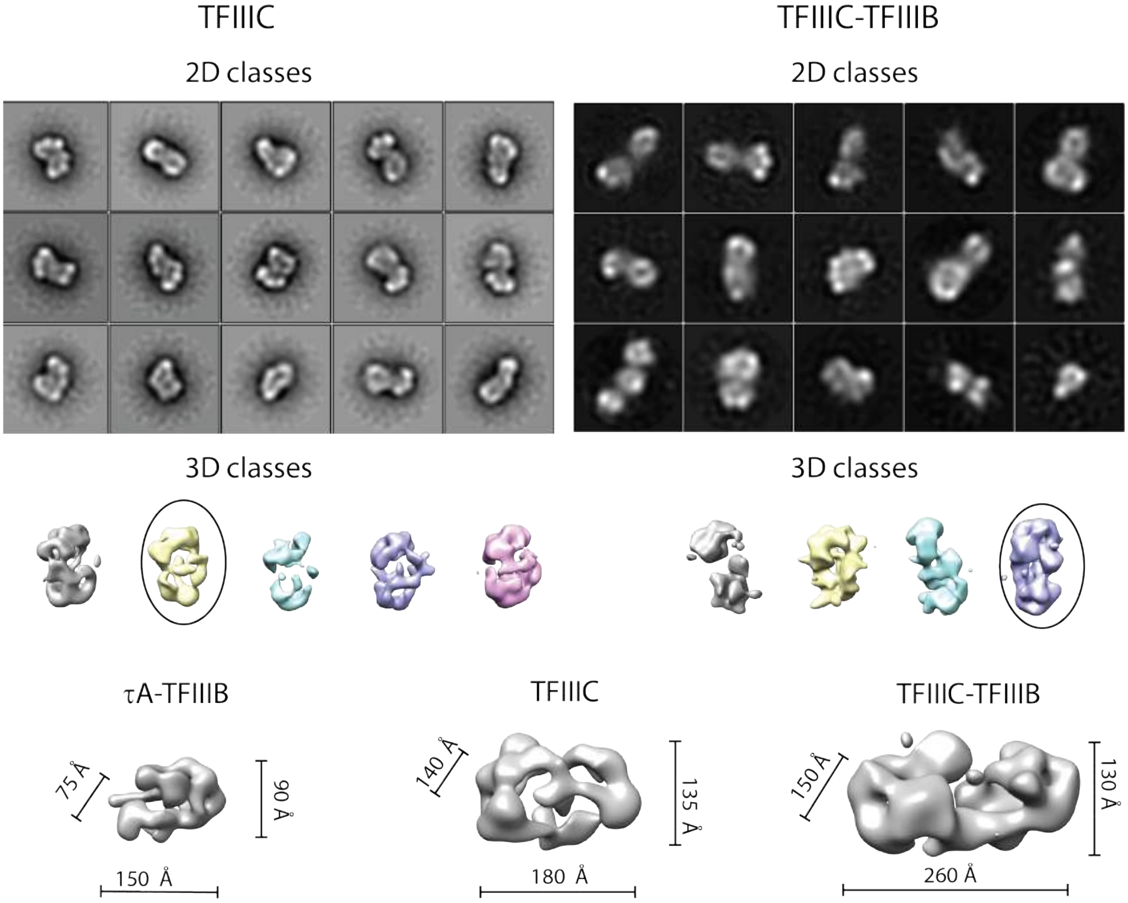


Figure 4.9: Comparison of negative stain reconstructions of TFIIIC and TFIIIC-TFIIIB. 2D classes and 3D classes are shown. Encircled classes in the middle panel are shown with corresponding dimensions in the panel below.

4.3.3 *τA is displaced from the Pol III-TFIIIB complex during promoter opening*

The recombinant production of τA also offers advantages for studying the structure of TFIIIC in complex with Pol III. First, the larger protein yields compared to holo-TFIIIC allows us to test more conditions. Second, using holo-TFIIIC required the use of long (~120 bp) DNA fragments. I suspect that unspecific binding of Pol III and TFIIIB to this DNA might be a cause of for the high degree of aggregation/oligomerization observed when Pol III is introduced. Using τA instead of TFIIIC allows us to reduce the length of the DNA to 56 bp. When the τA-TFIIIB-Pol III complex was analysed using SEC, it however became obvious that τA is displaced from TFIIIB and elutes separately from Pol III, despite the τA-TFIIIB complex being stable in absence of Pol III under the same conditions (Figure 4.10). This suggests that τA is displaced from the complex once DNA is melted, in agreement with the fact that the A-box sequence is covered by Pol III when mapped on the available structures.

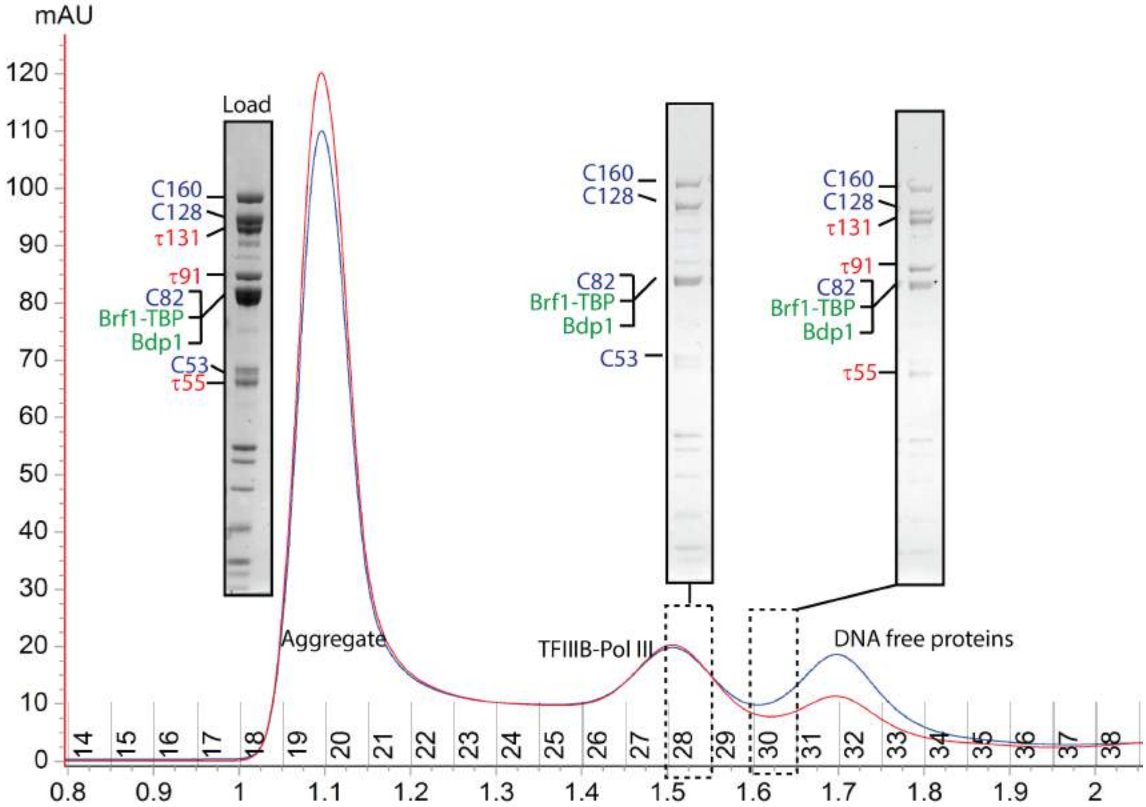


Figure 4.10: SEC analysis of a τA-TFIIB-Pol III complex. τA is only found in the the DNA-free fraction, whereas TFIIIB is bound to Pol III, indicating that Pol III displaces τA from TFIIIB. Pol III subunits are labelled blue, τA subunits are labelled red, TFIIIB subunits are labelled green.

This is supported by structural modelling. By superimposing TFIIIB in the TFIIB-τA model obtained from negative stain with TFIIIB in the structure of the pre-initiation complex containing closed DNA (PDB 6f44), I derived a model for the TFIIIC-containing PIC (Figure 4.11). This model contains only minor clashes between τA and TFIIIB-Pol III, and predicts an interaction of τ131 with the Pol III protrusion domain and the Pol III heterodimer-subcomplex, agreement with the interaction detected in yeast two-hyrid screens of τ131 and C57.

However, when the same complex is modelled with melted DNA (PDB 6f40), severe clashes are observed between the τ131 TPR and the C34-WH domains. This suggests that τA is displaced from Pol III concurrent with transcription initiation. Given that Pol III has to transcribe through the downstream sequence that is occupied by TFIIIC, this mechanism might ensure efficient transcription elongation.

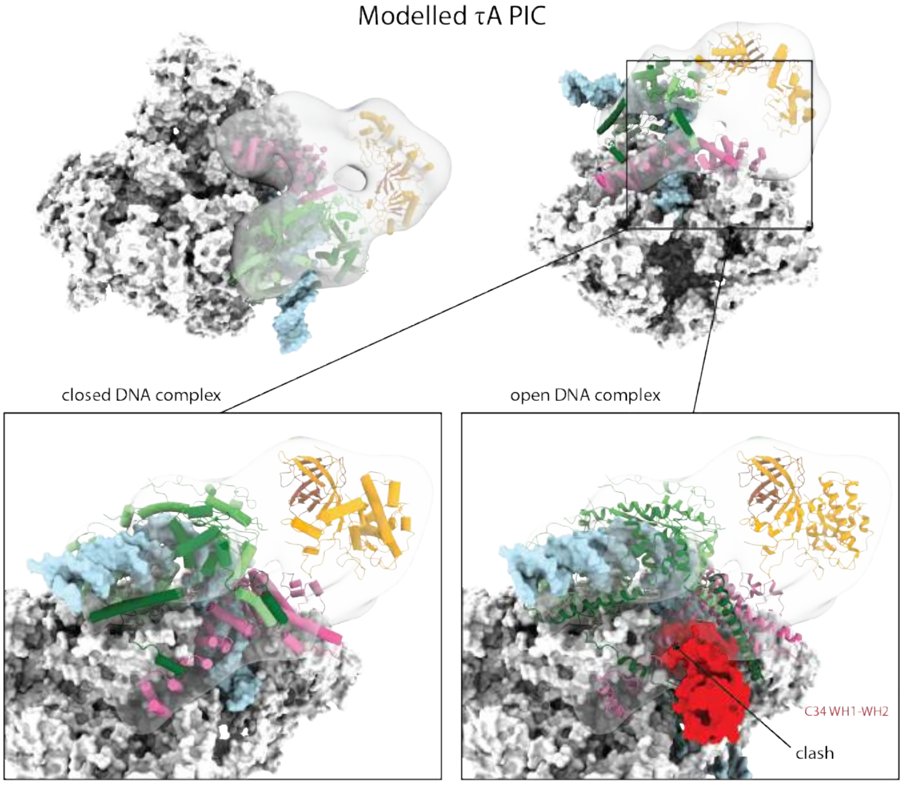
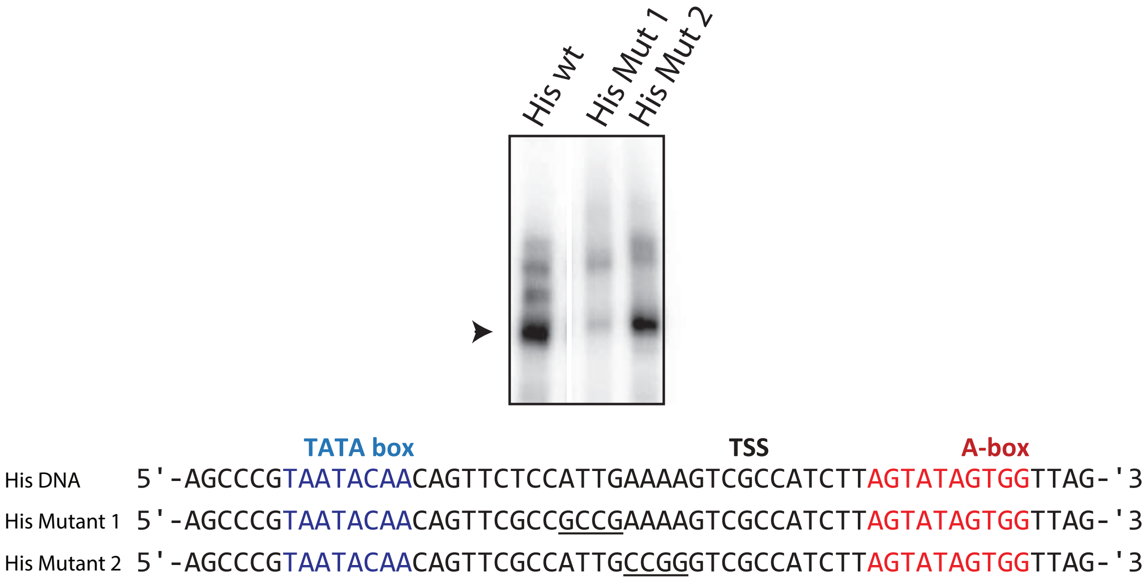


Figure 4.11: Model of the TFIIIC-PIC with negative stain density superimposed. While the closed DNA model appears plausible, the C34 WH domains clash with the τ131 TPR when DNA is melted.

Since the position of C34 in the promoter-melted state appears to displace τ131, we tested strategies designed to prevent promoter opening in the PIC (together with A.J.). First, I cloned, expressed and purified mutants in Brf1 and Bdp1 that were reported to prevent promoter opening, namely the Brf1-Δ(1-68) mutant and the Bdp1-Δ(355-372) mutant described earlier.

Second, I replaced nucleotides in the *His* promoter with GC stretches. This was motivated by a recent study of the Pol II system, which showed that stable DNA duplex require the factor TFIIH for promoter opening, whereas DNA duplex with a lower free energy can be opened without TFIIH [45].

The mutated promoter sequences were tested in *in vitro* transcription assays. This revealed that one of the mutant that had the GC stretches inserted at positions -11 to -8 relative to the TSS was very poorly transcribed, while expression from a promoter that had the GC stretches inserted from -7 to -4 was similiar to the wildtype promoter sequence (Figure 4.12).



His DNA	5'-AGCCCGTAATACAACAGTTCTCCATTGAAAAGTCGCCATCTTAGTATAGTGGTTAG-'3
His Mutant 1	5'-AGCCCGTAATACAACAGTTCGCCGCCGAAAAGTCGCCATCTTAGTATAGTGGTTAG-'3
His Mutant 2	5'-AGCCCGTAATACAACAGTTCGCCATTGCCGGGTCGCCATCTTAGTATAGTGGTTAG-'3

Figure 4.12: *In vitro* transcription of his-tRNA promoter mutants. Transcription from the His Mut1 promoter is severely impaired, whereas His Mut2 supports similar levels as the wild-type promoter. Cropped lanes from the same gel. Mutant sequences are given. Experiment performed by F.B.

Reconstitution experiments using these components are summarized in table 4.1. In brief, while some co-elution was detected under low salt conditions, this appeared to be due to peak broadening in low salt and no complex could be visualized by negative stain EM. This is exemplified in Figure 4.13.

Table 4.1: Summary of reconstitution experiments with mutant and wildtype TFIIIB and promoter DNA

Brf1-TBP	WT	Delta	Delta	Delta
Bdp1	WT	Delta	WT	Delta
DNA	WT	WT	WT	Mutant
observed complexes	stable Pol III-TFIIIB	reduced TFIIIB binding to Pol III	salt-sensitive TFIIIB binding to Pol III	some co-elution observed
	separate τA	separate τA-TFIIIB	sub-stochiometric τA, no complex observed in NS	mostly τA-TFIIIB co-elution

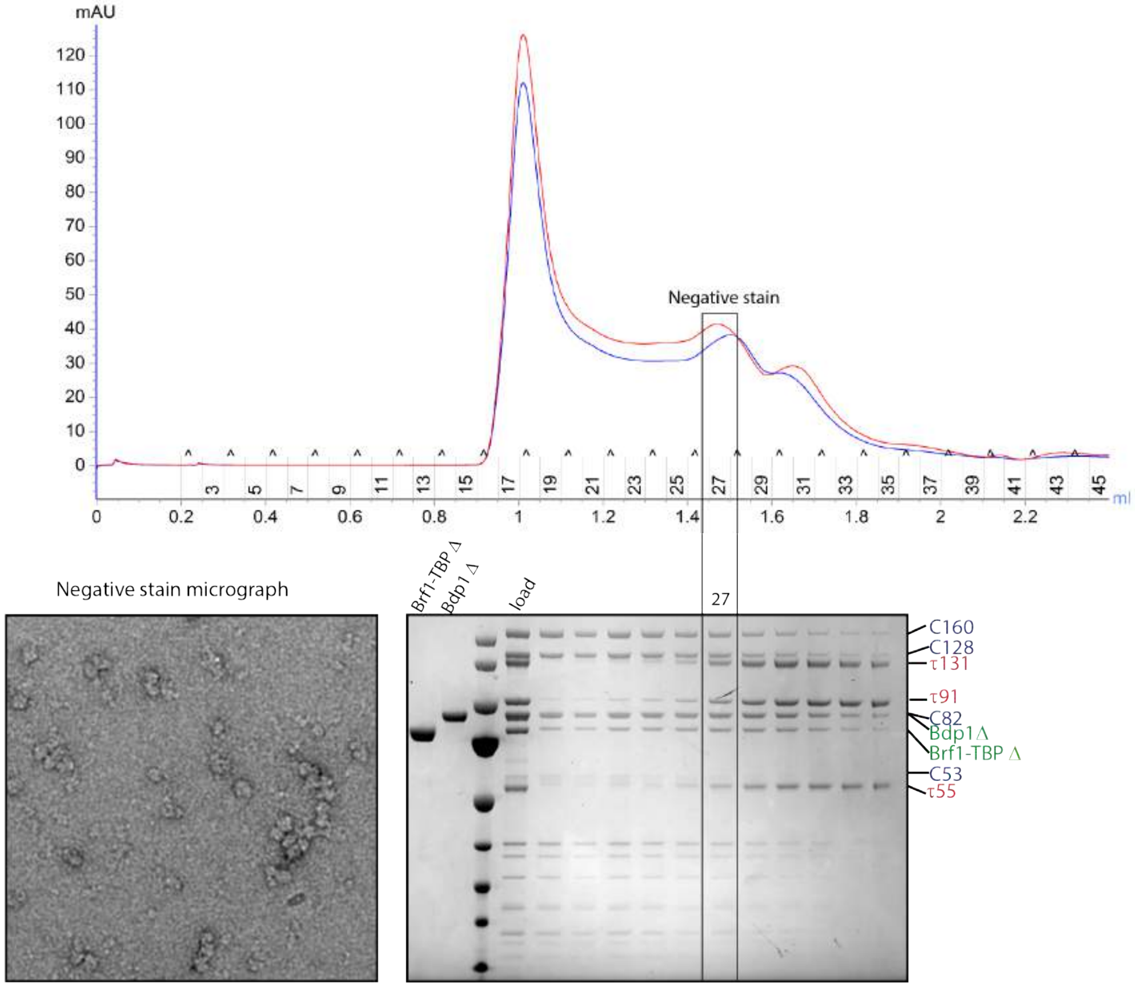


Figure 4.13: SEC and negative stain analysis of Pol III, Bdp1Δ, Brf1-TBPΔ and mutated *His* promoter. The fraction used for negative stain grid preparation is indicated. Although all subunits are visible in SDS-PAGE, partciles look heterogenous and aggregated in negative stain. Experiment performed by A.J.

Part III

DISCUSSION

5

DISCUSSION

5.1 THE ARCHITECTURE OF THE POL III PIC IS ADAPTED TO HIGH-FREQUENCY INITIATION

Pol III genes are very highly transcribed, with an estimated transcription rate of 2–4 transcripts per gene per second for the average tRNA gene [93, 145], and a new Pol III initiating on the same 5S rRNA gene every second [60]. The average density of Pol III on its target genes in yeast has been estimated to be one Pol III every 72 bp, but up to three Pol IIIs were observed by electron microscropy on the same 132 bp 5s rRNA gene [60]. In contrast, only 1% of yeast Pol II genes are estimated to have more then one engaged Pol II [156], although their average length is much greater than that of tRNA genes. Accordingly, the transcription rate of the most strongly expressed genes—the histones—has been estimated to be 206 transcripts per hour, or 0.06 per second [156]. The maximal initiation rate for Pol II genes has further been estimated at once every 5 seconds [44]. In the following section, I will draw comparisons between the molecular machineries underlying transcription initiation by Pol II and Pol III and speculate what enables the Pol III machinery to achieve such high initiation rates. Basic to this discussion is the architecture of the Pol III PIC described in the Results section. In brief, TFIIIB forms a highly stable promoter-DNA-binding core, and has two arms which provide a binding platform that stimulates the intrinsic transcription-factor-acticity of the Pol III heterotrimer.

5.1.1 *Comparison of the Pol III PIC architecture with Pol II and Pol I*

I compared the Pol III OC structure to the yeast Pol II OC (PDB 5fyw) to better understand the roles of general transcription factors and transcription factor-like subcomplexes in both systems (Figure 5.1). The overall topology of the PICs is highly similar in Pol III and Pol II, as both use an upstream promoter assembly centered on the TATA box that introduces a 90 ° kink, and a downstream assembly in which several WH domains position the upstream bubble edge along the cleft. In the Pol III PIC, the upstream assembly consists of the TFIIIB core, and in Pol II of TFIIB, TBP (or the much larger TFIID) and TFIIA. The downstream assembly in Pol III is provided by the heterotrimer where the transcription bubble is stabilized by C34 WH2 and the C82 cleft loop (Figures 2.13, 5.1). In the Pol II PIC, TFIIF and TFIIE occupy similar positions as C34 and C82, respectively. The 'E-wing' in TFIIE likely stabilizes the transcription bubble in a similar manner as the C82 cleft loop. Therefore, the overall architecture of the downstream assembly is conserved, further confirming the hypothesis that general transcription factors have been stably incorporated into Pol III during evolution [24, 201].

Moreover, our structures suggest that the "TFIIE-like" heterotrimer combines functions of both TFIIF and TFIIE, since the heterotrimer contributes equivalent DNA-binding WH-domains as TFIIE and TFIIF. However, similarity between the Pol II and Pol III PICs is restricted to a topological level, as the WH domains in both systems do not superimpose. In particular, the C34 WH2 binds more downstream compared to the corresponding TFIIF

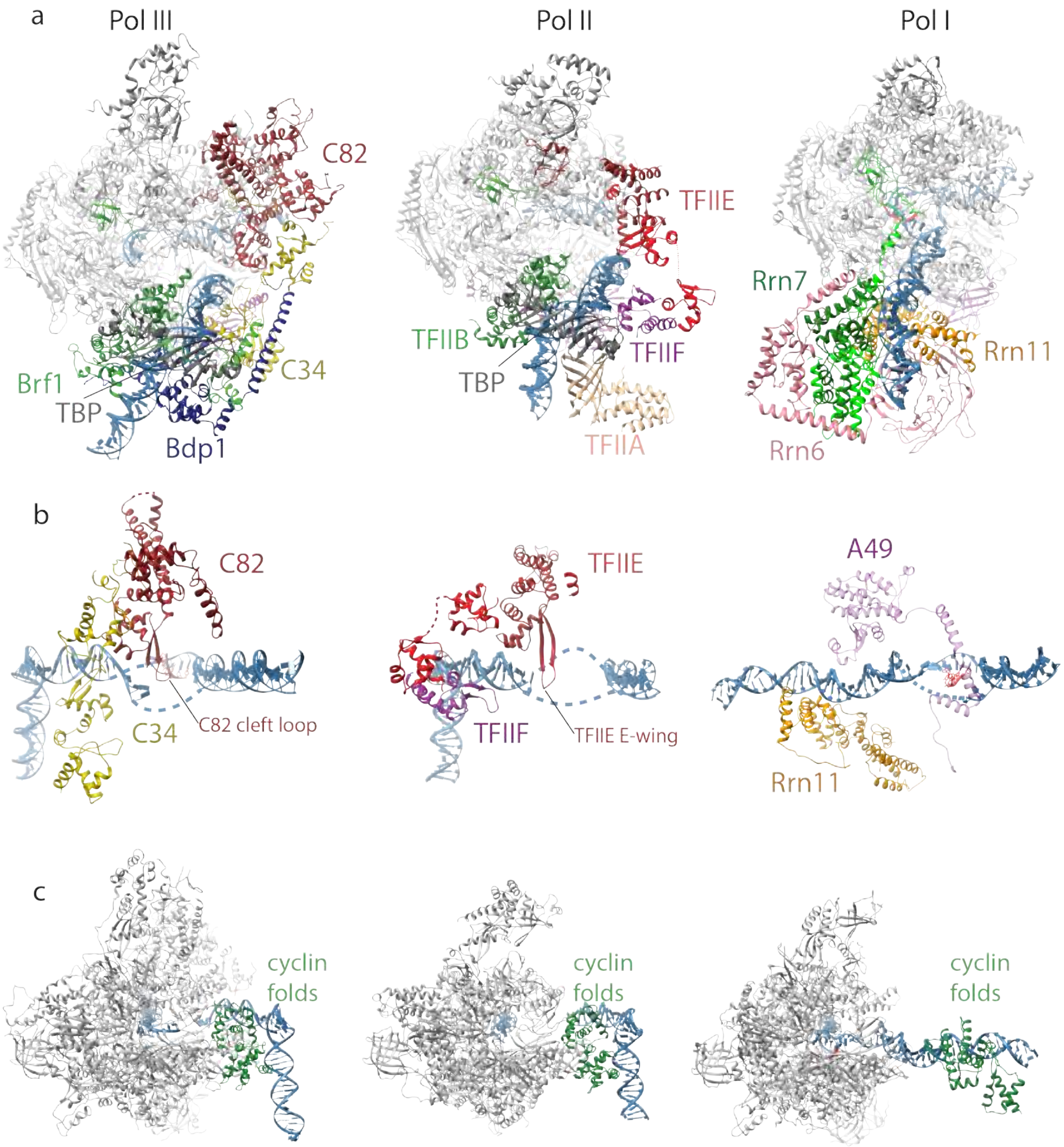

a
Pol III
Pol II
Pol I
C82
TFIIE
Rrn7
TFIIB
TFIIF
Rrn11
C34
Brf1
TBP
TBP
Bdp1
TFIIA
Rrn6
b
C82
TFIIE
A49
C34
C82 cleft loop
TFIIF
TFIIE E-wing
Rrn11
c
cyclin folds
cyclin folds
cyclin folds

Figure 5.1: Comparison of Pol I, II and III PICs. a, Ribbon diagrams of yeast Pol III, Pol II (PDB 5fyw) and Pol I (PDB 5w65) PICs. TFIIF and TFIIE occupy similar positions as C53-C37 heterodimer and C82-C34-C31 heterotrimer, respectively. The convex surface of TBP is closely contacted by Brf1 homology domain II in TFIIIB, but accessible in the Pol II PIC, whereas TBP is entirely absent from available Pol I PIC structures. This might explain the strict requirement of TBP in the Pol III system in contrast to Pol II and Pol I. b, Close up of the downstream promoter assembly, showing C82 WH3-WH4, C34 WH1-WH3 (left), the WH domains of TFIIF and TFIIE (middle) and the A49 tWH9 and TPR of Rrn11. While C82-C34 and TFIIF-TFIIE form structurally similar downstream promoter assemblies using WH domains and the C82 'cleft loop' and TFIIE 'E-wing' to contact the upstream bubble edge, Pol I CF forms a structurally different assembly where the A49 tWH does not contact the upstream bubble edge in the same way. c, Comparison of the cyclin folds in Brf1, TFIIB and Rrn7. The cyclin folds in Brf1 and TFIIB occupy similar positions and contact the polymerase wall, whereas the cyclin folds in Rrn7 do not. Figure taken from [203].

WH, and is thus likely to directly contribute to DNA melting by stabilizing the bubble.
I also notice that the position of the Bdp1 SANT domain in the PIC is similar to TFIIA in the Pol II PIC, and both Bdp1 and TFIIF exchange β-strands with the protrusion (Figure 5.2) , suggesting that Bdp1 combines functions of TFIIF and TFIIA.

Despite the similar overall architecture of the Pol II and Pol III PICs, the way the promoter assemblies are formed is different. Pol III only requires TFIIIB, which forms very stable complexes on promoters and stays bound after each initiation event [124], while both arms of TFIIIB provide a scaffold for the inbuilt TF-like subcomplexes of Pol III. This presumably underlies the ability of Pol III to achieve very high initiation frequencies. In contrast, the minimal Pol II PIC is transient and requires the factors TFIIB, TFIIE, TFIIF and TFIIH, of which TFIIB and pre-formed Pol II-TFIIF have to rebind for each transcription initiation event [43, 70, 162, 216]. This presumably allows for tighter control but results in a lower initiation frequency.

The architecture of the Pol I PIC that comprises the heterotrimeric core factor (CF) and Rrn3 is different from the Pol II and Pol III PICs, as in the Pol I PIC upstream DNA already deeply penetrates the cleft [52, 73, 174] (Figure 5.1). In addition, while the cyclin folds of Brf1 and TFIIB contact the polymerase wall, the cyclin folds of Rrn7 do not. The tWH domain of A49 might provide a 'downstream assembly' (Figure 5.1b) in Pol I, but it is disordered in most cryo-EM reconstructions and lacks elements equivalent to the C82 cleft loop or the TFIIE E-wing. Therefore, the promoter assembly in Pol I is structurally, and likely mechanistically, different from Pol II and Pol III.

5.2 MAF1 PREVENTS POL III TRANSCRIPTION INITIATION BY STERIC HINDRANCE

In this thesis, I present the first high-resolution structure of Maf1 bound to Pol III. The structure shows that Maf1 binds in a position that overlaps with promoter DNA and the Brf1 cyclin I in the pre-initiation complex. Residues in Pol III that form important interactions with DNA and TFIIIB in the pre-initiation complex are sequestered by Maf1, in particular a conserved tryptophane residue in the Pol III clamp helix forms an aromatic stacking with an invariantly conserved tryptophane in Maf1. Furthermore, in the Maf1-bound state, the C34-WH2 domain becomes ordered and binds Maf1. This binding appears to be dominated by salt bridges, as the shape and charge of two loops in Maf1 mimics the backbone of B-DNA. A comparison between yeast Maf1, human Maf1 and plant Maf1 shows that this acidic patch is conserved between the species, hinting at a conserved mechanism of repression (Figure 5.3).

The acidic C-terminus of Maf1 is disordered but projects towards the active site cleft where it might repel nucleic acids through electrostatic interactions.

The structure presented here explains why binding of Maf1 and Brf1-TBP are mutually exclusive [201]. In a previous low-resolution cryo-EM study, Maf1 was reported to bind Pol III at a different location, namely on top of the clamp. This had been validated by gold labelling of a his-tag on Maf1 and a his-tag on Pol III subunit C128 and distance measurements in 2D cryo-EM images [201]. However, the proposed binding site is located fairly close to our binding site, and the precision of the distance restraints from gold labels is not sufficient to discriminate against the two binding modes, especially considering that the first 38 amino acids of C128 are disordered in available cryo-EM structures and therefore likely flexible.

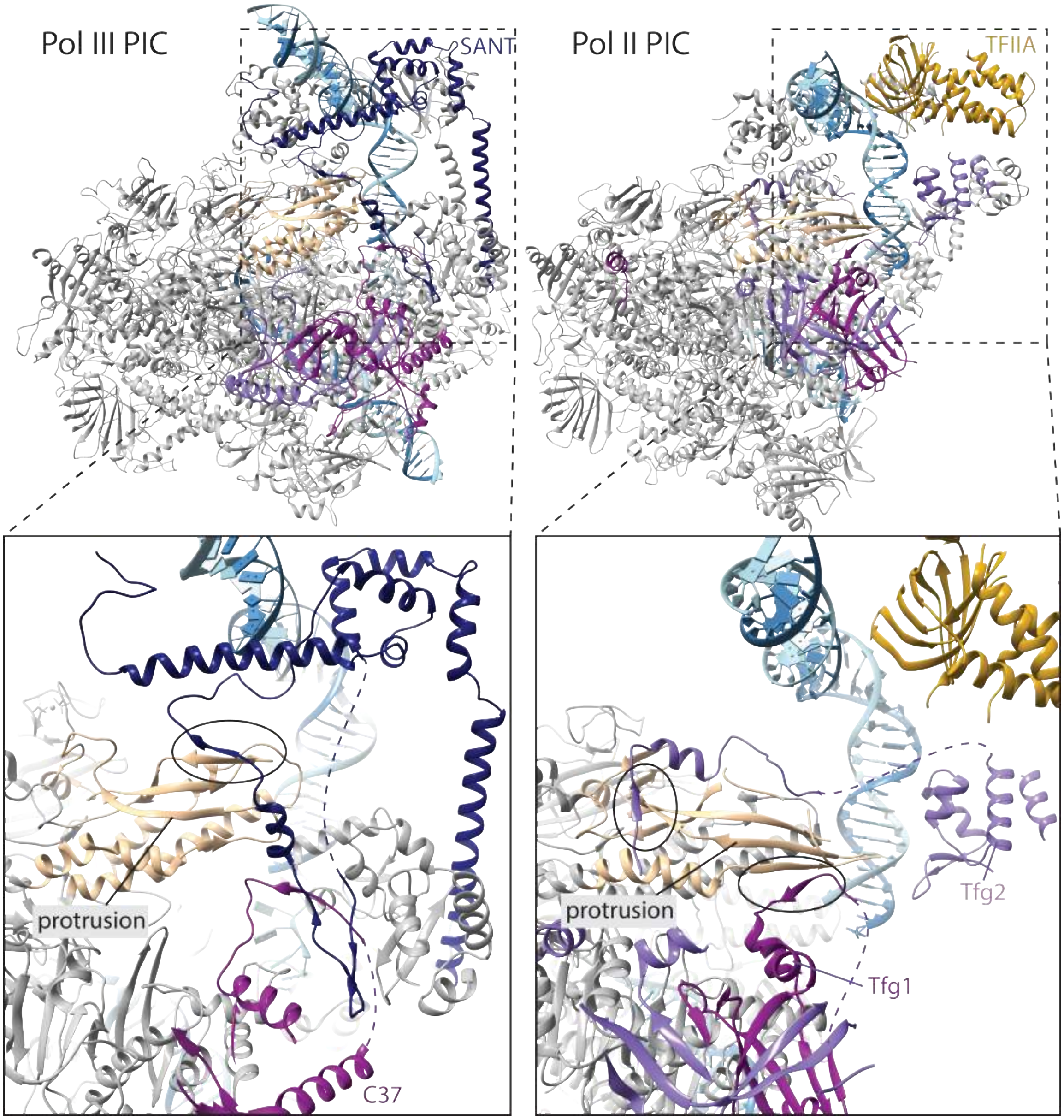

Pol III PIC
SANT
Pol II PIC
TFIIA
protrusion
C37
protrusion
Tfg2
Tfg1

Figure 5.2: The Bdp1 SANT domain is located at a similar position as TFIIA. Parts of Bdp1 resemble TFIIF subunit Tfg1, although no sequence similarity is detectable. Both interact with the Pol protrusion by adding a β-strand to the Pol protrusion (although at different ends of the protrusion β-sheet) and fold into a short helix along the face of the protrusion. The path of the C37 initiation/termination loop is also similar to Tfg1. The second subunit of TFIIF, Tfg2, also adds a β-strand to the Pol II protrusion. Brf1 and TBP were omitted for clarity. Figure taken from [203].

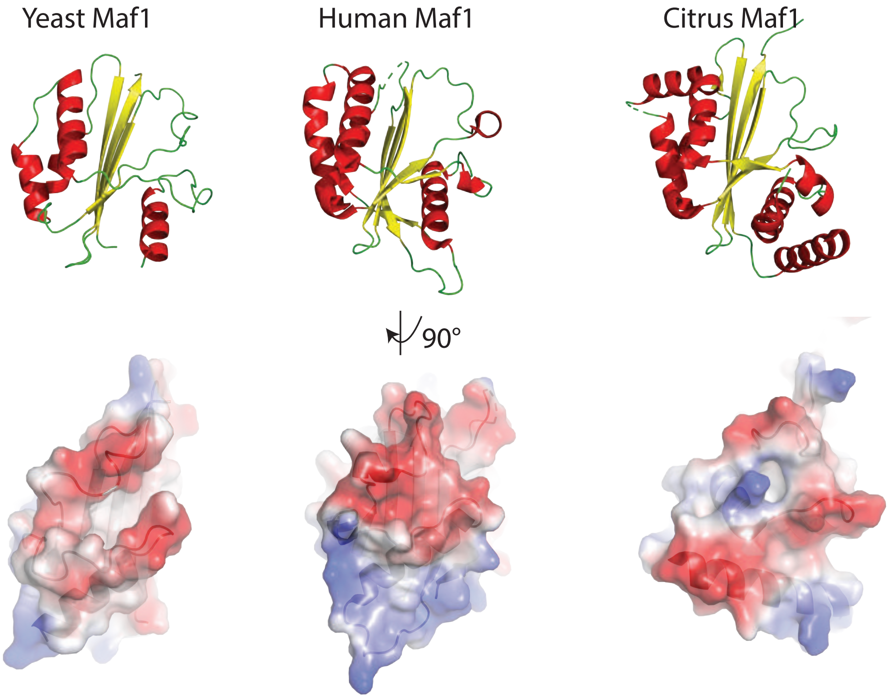


Figure 5.3: The Maf1 acidic patch is conserved between yeast, human (PDB 3nr5) and plant (PDB 5u4z). Top: Secondary strucure rendering. Bottom: electrostatic potentials mapped on the Maf1 surface.

Importantly, in the original paper the mechanism of repression that was suggested is an allosteric one. Comparison of cryo-EM structures from Maf1-Pol III complexes and Maf1-nucleic acid complexes led the authors to propose that Maf1 rearranges the hetrotrimer complex and thereby locks it in a conformation that is unable to bind TFIIIB. The structure presented here suggests direct competition for the same binding site.

While the improved resolution of our reconstruction gives confidence in the correctness of our data, further validation experiments are currently being pursued. Specifically, I designed mutations in Maf1 and Pol III that I expect to disrupt the binding of Maf1 to Pol III (shown in Figure 5.4). The Maf1 mutants have been cloned, expressed and purified and will be tested in binding Pol III and recombinant C34(1-156), and in a promoter-dependent *in vitro* transcription assay.

Pol III mutants were generated in the lab of Ian Willis and will be tested in an *in vivo* assay.

5.3 TOWARDS A CRYO-EM STRUCTURE OF TFIIIC

TFIIIC has been a focus of the research in the Müller lab for many years, and this resulted in a number of crystal structures of individual domains, subcomplexes and a coarse-grain architecture based on XL-MS restraints. However, structures of holo-TFIIIC or larger submodules are missing. In part, this was hindered by the low amounts of TFIIIC that can be purified from natural sources and the lack of a recombinant expression system. Here, I describe recombinant expression of holo-TFIIIC and the τA module. The latter is capable of stimulating basic levels of transcription. Production of recombinant τA also enabled the determination of the first interpretable negative stain structure of τA bound to TFIIIB. In this structure, TFIIIB is bound to the N-terminal TPR array of τ131, in agreement with

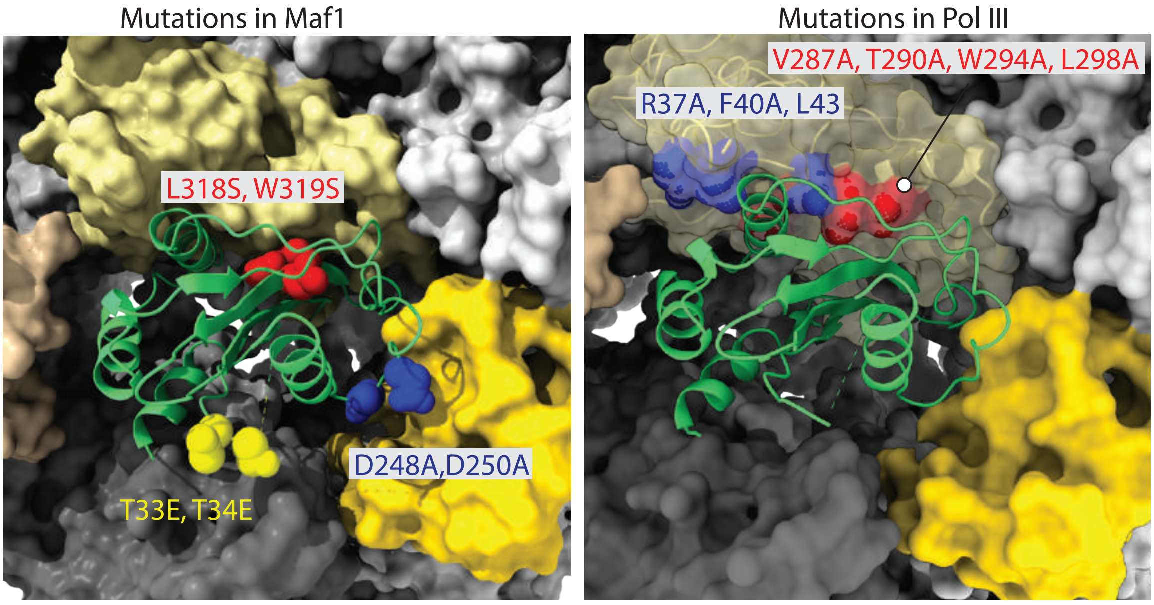


Figure 5.4: Positions of engineered mutations in Maf1 and Pol III are indicated

biochemical data. The map shows a potential route for promoter DNA which connects from TBP to the other end of the molecule. The TPR array thus functions as a molecular ruler, which positions TFIIIB at a (relatively) fixed position upstream of the TSS, where it can then bind to DNA sequence independently of a TATA box.

This is reminiscent of a recent cryo-EM structure of TFIID, the large transcription factor of the Pol II system which binds downstream of the TSS to its promoter elements and enables binding of TBP upstream of TSS at genes lacking a TATA box [132]. Like TFIIIC, TFIID is a large, flexible transcription factor, and its structure was only determined at low resolution (~10 Å) by collection large amounts of cryo-EM data.

Initial attempts to obtain interpretable cryo-EM maps of TFIIIC (work by former postdoc H.K.) have been fruitless and have therefor been put to rest. Given the example of TFIID, it is however perhaps worth to invest further effort into this project. A structure of holo-TFIIIC would be of great interest to the Pol III field, as well as the community interested in the extra-transcriptional roles of TFIIIC.

The negative stain and initial cryo-EM maps of τA TFIIIB have already provided some insights about how TFIIIB is assembled upstream of the TSS, however, high resolution information is required to confidently place the τ95 and τ55 structures and to build a model of the C-terminal TPR array of τ131. Moreover, no density for DNA is observed, which is an inherent feature of negative stain reconstructions of protein-DNA complexes. Therefore, the sample is currently being optimized for cryo-EM.

This provides a number of challenges. First, the complex is rather small (345 kDA, including some most likely disordered/flexible regions), limiting image contrast especially on carbon grids. Therefore, grids without carbon support have been tested, but no particles could be observed so far.

A second challenge is the instability of the complex under EM conditions. Both in negative stain (data not shown) and cryo-EM, particle dissociation was observed when no crosslinking was used. On the other hand, crosslinked sample showed a lot of large particles presumably resulting from inter-particle crosslinks, even under mild crosslinking conditions. Presumably for this reason only a rather small fraction of the particles in the negative stain dataset (10-20%) gave rise to the negative stain reconstruction.

To make the complex amenable for cryo-EM, it might be beneficial to use EM grids covered with a monoloayer of graphene oxide, which provides a support for particles to adsorb to but produces a much better contrast compared to carbon. Another option is the use of detergents, which have been shown to prevent the adsorbtion of particles to the air-water interface and thereby might counteract some of the deleterious effects of thin film-formation. Initial screening micrographs show particles in grids prepared with 4 mM CHAPSO, however particles were very sparse. Therefor, a much higher concentration might be required.

Another project that is currently ongoing is the EM characterization of the τB module. This is being pursued by Anna Jungblut. Initial insect cell expression constructs have been tested and expression was detected, but purification has been hampered by poor binding to StrepTactin beads directly from a lysate, thus requiring cloning of additional affinity tags.

5.4 TFIIIC FUNCTIONS AS AN ASSEMBLY FACTOR

In this thesis I demonstrated that the τA module is displaced from TFIIIB when Pol III is added. In line with this, negative stain EM reconstructions of Pol III incubated with TFIIIC-TFIIIB revealed the presence of TFIIIB-Pol III complexes, but no TFIIIC-TFIIIB-Pol III complexes were detected. Structural modelling of the τA-TFIIIB-Pol III complex revealed that severe clashes with the C34 WH domains occur in the promoter-melted state, whereas binding of τA appears to be possible in the the closed-DNA complex where the C34 WH domains are disordered.

Efforts to stabilize a τA-contianing PIC were made by using mutants in TFIIIB that are deficient in promoter opening, or by using DNA scaffolds with increased stability, however these appear to affect the stability of the TFIIIB-Pol III complex as well and no convincing evidence for formation of stable complex containing τA and Pol III was found.

Since TFIIIC binds downstream of the TSS, displacement of τA from its promoter sequence by Pol III might be a mechanism to achieve efficient transition into the elongation phase. In the context of holo-TFIIIC, the τB module is likely still bound to the B-box at this stage. Therefore, a hand-over mechanism, in which TFIIIC is bound to transcribed Pol III genes alternating between a τA-bound and a τB-bound state, might occur. However, evidence from *in vivo* chromatin occupancy studies showed a decrease of TFIIIC binding to Pol III during active transcription and an increase when transcription is repressed [33, 75]. This argues in favour of a displacement of TFIIIC by Pol III.

Based on the presented structural modelling of the τA PIC, another potential way to stabilize this complex would be by deleting the C34 WH1-WH2 domains in Pol III, as the clash between these and τ131 likely results in the disassembly of τA. This could be achieved by engineering a TEV site in the linker between C34 WH2 and WH3[1], as yeast strains lacking the C34 WH1-WH2 are unlikely to be viable. Although the physiological relevance of such a structure would be debatable, it might be more feasible to determine an EM map of such a complex due to its large size and the strong signal that Pol III provides during image alignment.

---

[1] residues 158-165 are disordered in all available structures and no seconday structure elements are predicted. This might be a suitable region for insertion of a TEV site

Part IV

MATERIAL AND METHODS

6

MATERIAL AND METHODS

6.1 CLONING OF RECOMBINANT TFIIIC

When the project was initiated, constructs carrying insect cell codon-optimized genes were already present in the lab, but some were controlled by a pH promoter and some by a p10 promoter. Therefore, all subunits were first cloned into the pAceBac vector using Gibson assembly [63]. During this step, an N-terminal His-TEV tag was added to τ91. Using restriction free-cloning [200], C-terminal TEV-cleavable Strep-tags were added to the genes of τ60 and τ138.

After all six subunits including the correct tags were cloned into the pAceBac vector, subunits were assembled using the 'homing endonuclease' strategy [10]. In brief, the acceptor vector (starting from pAceBac- τ138) was linearized with BstXI and dephosphorylated with alkaline phosphatase, and the insert was cut from the donor vector with BstXI and ICeuI and purified via gel extraction following standard protocols. Vector and insert were then ligated using T4 DNA ligase, yielding the acceptor vector for the next cycle of cloning. The tags on the final construct are summarized in table 6.1.

Later on, a new expression construct was made were the Strep-tag of τ60 and the His-tag of τ91 were removed, in order to simplify the purification procedure (performed by A.J.). This greatly improved sample homogeneity after the Strep-tactin affinity purification step.

Table 6.1: Expression construct for recombinant TFIIIC production in insect cells

τA	
τ131	No tag
τ95	N-His-TEV
τ55	No Tag
τB	
τ138	C-TEV-TwinStrep
τ91	N-His-TEV
τ60	C-TEV-TwinStrep

6.2 PROTEIN EXPRESSION AND PURIFICATION

6.2.1 *Pol III*

Endogenous Pol III was purified as described [148] with minor modifications. The glycerol content in the lysis buffer was reduced to 20 %, and the Heparin column was washed overnight with HepA instead of lysis buffer. Finally, the protein was buffer exchanged into a buffer containing Li_2SO4 or K_2SO4 instead of $(NH_4)_2SO4$ for protein crosslinking.

6.2.2 *Brf1-TBP*

The Brf1-TBP plasmid[134] was transformed into BL21 Star (DE3) pRARE *E. coli* cells. Expression cultures were grown at 37 °C in TB medium to an optical density at 600 nm (OD) of ~1.0, cooled down (for 1 h at 4 °C) and induced with 50 µM IPTG overnight at 16 °C. Cells were pelleted for 5 min at 12,000 g and re-suspended in 3 mL lysis buffer (1M NaCl, 50 mM Tris pH 7.5, 2 mM β-mercaptoethanol (BME), 20% glycerol, 10 µg/mL DNase I, 1 x protease inhibitors (SIGMAFAST protease inhibitor cocktail EDTA free), 30 mM imidazole, 2 mM $MgCl_2$) per gram of pellet. Cells were lysed in an Emulsiflex-C3 homogenizer and the lysate cleared by centrifugation for 1 h at 30,000 g. The supernatant was incubated with 5 mL Ni-NTA resin (Qiagen) for 2 h. Beads were recovered and washed with 100 mL His-A buffer (1 M NaCl, 50 mM Tris pH 7.5, 2 mM BME, 5% glycerol, 30 mM imidazole) and 50 mL His-A low salt (His-A but with 150 mM NaCl) and eluted with 50 mL His-B (50 mM Tris pH 7.5, 200 mM NaCl, 2 mM BME, 5% glycerol, 300 mM imidazole). The eluate was loaded on a 5 mL HiTrap Heparin column (GE healthcare) pre-equilibrated in HepA buffer (like HisB but without imidazole). The column was washed with 6 column volumes (CV) containing 30% HepB (like HepA but with 1 M NaCl) and eluted with a linear gradient from 30% HepB to 70% HepB over 20 CV. Brf1-TBP eluted at 600 mM NaCl. Peak fractions were concentrated and applied to a HiLoad 16/600 Superdex 200 size exclusion column equilibrated in 300 mM NaCl, 25 mM HEPES pH 7.5, 5 mM DTT and 5 % glycerol. Purified Brf1-TBP was concentrated to ~6 mg/mL and flash frozen in liquid nitrogen and stored at -80 °C.

6.2.3 *Bdp1*

The Bdp1 plasmid [134] was transformed in BL21 Star (DE3) pRARE *E. coli* cells and grown in TB medium to an OD of ~1.0, cooled down and induced with 100 µM IPTG overnight at 18 °C. Cell harvesting, lysis and Ni-NTA chromatography were performed as for Brf1-TBP. Eluted proteins were loaded on a 5 mL Heparin column pre-equilibrated in HepA. The column was washed with 6 CV of 20% HepB and eluted with a gradient from 20% HepB to 70% HepB over 30 CV. Peak fractions (~520 mM NaCl) were digested with TEV protease overnight at 4 °C and incubated with 3 mL Ni-NTA for 1 h. The column was washed with 10 mL HepA containing 100 mM imidazole to recover cleaved Bdp1, which bound to Ni-NTA unspecifically. Bdp1 was finally purified by size exclusion chromatography as for Brf1-TBP (but in a buffer containing 150 mM NaCl) and concentrated to ~6 mg/mL. Aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

An alternative protocol with higher yields was also used. For this, a construct carrying an N-terminal-glutathione-hexahistidin tag followed by a TEV cleavage site was expressed from a pETM29 vector. Expression and initial purification (until the Ni-NTA elution) were performed as described above. The eluate was incubated with 1 mg of TEV protease overnight at 4 °C. The cleaved protein was purified over Heparin and Superdex 200 columns as described above.

6.2.4 *Maf1*

Maf1 constructs were expressed from pET30(a) (full-length Maf1) or pETM11(Maf1 i.d .) vectors in BL21 Star (DE3) pRARE *E. coli* cells in TB medium. Expression was induced with 500 µM IPTG at 25 °C overnight. Cells were resuspended in His-A buffer but with 500 mM NaCl and purified over Ni-NTA as described for Brf1-TBP. For full-length Maf1, Ni-NTA beads were further washed with 30 mL chaperone release buffer (HisA low salt with 10 mM ATP and 20 mM $MgCl_2$ and 0.2 mg/ml heat denatured *E. coli* lysate, pre-warmed to 37 °C). For TEV-cleavable constructs, the eluate was supplemented with 1 mg of TEV protease and dialysed against 2 L of TEV-cleavage buffer (100 mM NaCl, 5 % glycerol, 1 mM DTT, 50 mM Tris pH 7.5) overnight. Cleaved protein was passed over 5 mL of Ni-NTA beads pre-equilibrated in TEV cleavage buffer. Full-length Maf1 and Maf1 i.d. were further purified over a MonoQ HiTrap column and eluted with gradient from 100 mM NaCl to 500 mM NaCL over 20 CVs. Maf1 i.d., which was very strongly expressed, sometimes eluted as two peaks due to a fraction of dimers present under high protein concentrations (~5-10 mg/mL, as estimated from dynamic light scattering measurements). Maf1 i.d. was buffer exchanged in a spin concentrator to 100 mM NaCl, 1 mM TCEP, 15 mM HEPES. Full-length Maf1 was further purified over a Superdex 200 SEC column equilibrated in 150 mM NaCl, 2 mM DTT, 5% glycerol, 15 mM HEPES pH 7.5.

6.2.5 *TFIIIA*

TFIIIA was initially expressed as an N-terminal hexa-his-TEV protein, but co-purified with C-terminal truncation products. Therefore, an additional TEV-cleavable C-terminal Strep tag was cloned into the construct.

TFIIIA was expressed from a pRSF-Duet vector as a codon-optimized gene in BL21(DE)3 *E. coli* cells. Expression was induced at an OD of 0.8 with 250 mM ITPG and 100 µM $ZnCl_2$ were added. Expression was done overnight at 18 °C. Cells were lysed in His-A supplemented with 100 µM $ZnCl_2$. Lysates were prepared and cleared as described for Brf1-TBP and TFIIIA was bound to 5 mL Ni-NTA. Beads were washed with 100 mL lysis buffer and 20 mL chaperone release buffer (Lysis buffer plus 5 mM ATP plus 20 mM $MgCl_2$, pH adjusted to 7.5 mM with additional Tris buffer). Proteins were eluted in HisB supplemented with 1 tablet of protease inhibitors (SIGMAFAST protease inhibitor cocktail EDTA free). TFIIIA was further purified over 5 mL StrepTactin resin pre-equilibrated in Strep-A (500 mM NaCl, 50 mM Tris pH 7.5, 2 mM DTT), washed with 30 mL Strep-A and eluted with Strep-A supplemented with 5 mM desthiobiotin. Eluted protein was cleaved with 1 mg of TEV overnight and the cleaved protein was passed over 2 mL Ni-NTA pre-equilibrated in StrepA supplemented with 75 mM imidazole. Finally, TFIIIA was purified over a Superose 200 SEC column pre-equilibrated in 500 mM NaCl, 20 mM HEPES pH 7.5 and 5 mM DTT.

6.2.6 *Recombinant TFIIIC*

Expression of TFIIIC was achieved by infecting High Five insect cells at a density of 1 million cells per mL culture with virus containing a bacmid encoding for all six subunits of TFIIIC at a dilution of 1 to 2000. Cells were harvested after 72 hours by centrifugation at 800g for 15 minutes, washed in cold phosphate buffered saline (PBS), centrifuged again and stored at -80 °C.

Cells were resuspended in 50 mL lysis buffer per liter of culture (lysis buffer: 500 mM (NH4SO4)2 , 50 mM Tris pH 7.5, 10 % glycerol, 30 mM imidazole, 2 mM BME, 5 µL Benzonase and 500 µL DNase I) and lysed by sonification (15 minutes total pulse time, 40% amplitude, 10 seconds on, 10 seconds off). The lysate was cleared by centrifugation for 1 h at 235 000 g and the supernatant filtered through 1.2 µm pores. The filtrate was loaded over a 5 mL Ni-NTA FF column at 1 mL/min using a sample pump, washed with 50 mL lysis buffer and eluted with 25 mL elution buffer (50 mM Tris pH 7.5, 200 mM NaCl, 5 mM BME, 300 mM imidazole). Peak fractions were supplemented with 1 protease inhibitor tablet and passed over 5 mL Strep-Tactin resin pre-equilibrated in Strep-A (150 mM NaCl, 50 mM Tris, 5 mM BME) at room temperature. Beads were washed with 25 mL Strep-A and eluted with Strep-B (Strep-A plus 5 mM desthiobiotin). At this stage, subunits τ91 and τ60 were present in excess of holo-TFIIIC and removed by Heparin chromatography followed by SEC over a Superose 6 column in 150 mM NaCl, 20 mM HEPES, 5 mM DTT.

6.2.7 *Recombinant τA*

τA was expressed from a bacmid carrying the τA genes shown in table 6.1. 3.6L of High five cells were infected with virus at a ratio of 1:500. Pellets were harvested and washed as described for TFIIIC and resuspended in 200 mL lysis buffer (500 mM NaCl, 10% glycerol, 50 mM Tris, 4 mM BME, 30 mM imidazole, 2 mM $MgCl_2$) and supplemented with 3 µL benzonase and 200 µL DNase (10 mg/mL). Cells were lysed by sonication, and cleared and loaded on a 5 ml Ni-NTA FF column as described for TFIIIC. Most of τA eluted in a wash step with 10% HisB (57 mM imidazole). This fraction was supplemented with 1 mg of TEV protease and dialysed against 2L of dialysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 2 mM DTT) overnight. The dialysed protein was further purified over Heparin, S200 and a MonoQ 5/50 GL column. For the last step, Buffer A contained 100 mM NaCl, 5% glycerol, 20 mM HEPES pH 7.5 and 5 mM DTT, and buffer B contained the same but 1M NaCl. Sample was eluted with a gradient from 0% B to 40% B over 40 mL. An excess of τ95-τ55 eluted at a NaCl concentration of ~290 mM NaCl (21%B) and τA eluted at ~356 mM NaCl (28.5% B). For future purifications it is advised that the Heparin and SEC steps are skipped, as the MonoQ 5/50 GL column provided the best separation by far. It might be necessary to dilute peak fractions of τA after the first round and re-apply them to the column for a second chromatographic separation to achieve the desired purity.

MonoQ-purified τA was buffer exchanged to 150 mM NaCl, 20 mM HEPES pH 7.5, 5 mM DTT in a spin concentrator and concentrated to 7 mg/mL before flash-freezing in nitrogen.

6.3 structure determination of the pol iii-tfiiib complex

6.3.1 *DNA oligo-nucleotides*

For the preparation of pre-initiation complexes, 81-nt long DNA scaffolds based on the U6 gene were used. The sequence contains the U6 promoter from -49 to +31 and, for the DNA-RNA hybrid, a 15-nt mismatch from -6 to +8 (template strand: 5'-CCAAATGTCCACGAAGGGTTACTTCGGCAACACATAGTTGCGAAAAAACA-

TTTATTTATAGTAGCCGAAAATAGTGGACG-3, non-template strand 5′-CGTCCACTATTTTCGGCTACTATAAATAAATGTTTTTTTCGCAGTCTATGCGGTTAACAGTAACCCTTCGTGGACATTTGG-3′, RNA 5′-GUUGCC-3′). For the DNA-RNA scaffold, we scrambled two positions in the template strand that base pair with the RNA (wild type sequence in template strand relative: +1-CAAGCG-+6, scrambled sequence: +1-CAACGG+6) in order to remove an alternative complementary site where the RNA could bind. The closed DNA scaffold contains the wild-type sequence (template strand: 5′-CCAAATGTCCACGAAGGGTTACTTCGCGAACACATAGTTGCGAAAAAAACATTTATTTATAGTAGCCGAAAATAGTGGACG-3′). Duplex DNA was generated by mixing single-stranded DNA oligo-nucleotides in H2O and heating to 95 °C for 10 min. The reaction was cooled down to 20 °C at a rate of 1.5 °C per min. For the DNA-RNA hybrid, RNA was added to the DNA duplex and heated to 40 °C for 10 min and then cooled down to 4 °C.

6.3.2 *Assembly of pre-initiation complexes*

For *in vitro* reconstitution of the Pol III-PIC we resorted to a fusion construct of Brf1 and TBP (Brf1-TBP) that has been shown to substitute Brf1 and TBP function *in vitro* and *in vivo* [108]. For the ITC, we pre-incubated Pol III with the ITC scaffold, which ensures positioning of Pol III with the correct polarity on the transcription bubble due to binding of the DNA-RNA hybrid in the active site. 300 µg of Pol III were incubated with a 1.1x excess of DNA-RNA scaffold for 20 min and Brf1-TBP and Bdp1 were added at a 3x excess for 1 h on ice. Samples were diluted to 0.5 mg/mL (calculated for Pol III) in crosslinking buffer (100 mM Li_2SO4, 15 mM HEPES pH 7.5, 10 mM DTT, 5 mM $MgCl_2$, 0.05% glutaraldehyde) and crosslinked on ice for 30 min. Crosslinking was quenched by addition of 40 mM Tris pH 7.5.

For the closed complex, the order was reversed and Brf1-TBP and Bdp1 were incubated with DNA for 20 min to position TFIIIB on the TATA box before Pol III was added. Crosslinked samples were concentrated in spin concentrators (Amicon Ultra, 500 uL, 30 K cutoff) and applied to a Superose 6 INCREASE 3.2/300 column equilibrated in EM buffer (150 mM Li_2SO4, 15 mM HEPES pH 7.5, 10 mM DTT, 5 mM $MgCl_2$). 50 µL fractions were collected and the peak fraction was used for cryo-EM grid preparation.

6.3.3 *Electron microscopy and data processing*

Cryo-grids were prepared with a Vitrobot IV set to 100% humidity and 4 °C. Quantifoil 200 mesh Cu 2/1 grids were glow discharged in a Pelco EasyGlow glow discharger and 2.5 µL of sample was applied (blotting parameters: wait time 15 s, blot force 4, blot time 4 s) and plunge-frozen in liquid ethane. Micrographs were acquired on a Titan Krios operated at 300 keV equipped with a Gatan Quantum energy filter and a K2 Summit direct detector. The detector was operated in super resolution mode at 105,000 magnification and a calibrated physical pixel size of 1.35 Å. Data collection parameters and dataset sizes are shown in 6.2.

For all datasets, frame alignment and dose weighting were performed with MotionCor2 [219] and contrast transfer function (CTF) parameters estimated with Gctf [217]. Particle picking and classification were performed with RELION 2.0 [117].

For the ITC dataset 773k auto-picked particles were extracted and binned 4 times to reduce computational costs. Particles were subjected to 3D classification using a 60 Å low-pass filtered model of apo Pol III (PDB 5fj9) without prior 2D classification. The major of the 4 classes contained 60% or 464k particles and showed clear density corresponding to TFIIIB and promoter DNA. Particles of that class were unbinned, extracted in 300 pixel boxes and refined. The resulting volumes showed clearly defined secondary structure elements for TFIIIB, but at a lower threshold compared to the Pol III core. Hence, I performed masked classification using a mask that covers TFIIIB ('TFIIIB mask 1'), upstream DNA and C34 WH1 and WH2. Out of the three classes, a smaller class with 79k particles (10.2% of autopicked particles) showed very well defined TFIIIB density at the same threshold as the Pol III core. The class had clear density for downstream DNA, but at a lower threshold then upstream DNA. Therefore I performed classification using a mask on downstream DNA. This yielded the OCΔdownstream-1 class (38k particles, 4.9%) and the ITC class (29k particles, 3.8%) maps.

We collected two datasets for the closed complex. Both were initially processed separately as the ITC dataset until the first masked classification using TFIIIB mask 1. I noticed that a minor class in both datasets had reduced density for C34 WH1 and WH2 and a different path of the upstream DNA. I pooled all particles that had strong TFIIIB density of both CC datasets (particle set "CC joined"), which contained 226k particles and was refined to 3.4 Å ("Pol III PIC joined" map). I classified the CC joined particle set with a global mask (250 Å diameter) and noticed that one class showed stronger DNA density in the cleft as well as additional density adjacent to Pol III. Reclassification of these particles with a 400 Å mask revealed a subset of dimers (3% of CC joined); in the dimers, two Pol III molecules bind to the same DNA molecule with their cleft. I excluded dimers and pooled all other classes, which I classified using a mask that covered C34 WH1 and WH2. From this I obtained two classes (142k particles, 67% of CC joined) that exhibited upstream DNA and C34 density as in the ITC/OC1 maps that were pooled and refined to 3.5 Å (not shown), and a minor CC class that had very weak C34 WH1 and WH2 density and showed the path of the upstream DNA (68k particles, 30% of CC joined). The CC class was further classified using a mask covering TFIIIB and upstream DNA as seen in the CC as well as C34 WH1 and WH2, which yielded the CC2 (34k particles, 15% of CC joined) and the CC1 reconstructions (19k particles, 8.2% of CC joined).

I further classified the 3.5 Å OC class using a mask on downstream DNA, which yielded the OC (62k particles, 29% of CC joined) and a map with disordered downstream DNA OCΔdownstream-2 (79k particles, 35.9% of CC joined).

6.3.4 *Model building*

I constructed an initial model by combining the structures of elongating Pol III (PDB 5fj8) and the Brf1-TBP core crystal structure (PDB 1ngm), encompassing TBP residues 61-240 and Brf1 residues 437-506. We next used Phyre2 [111] to calculated homology models of Brf1 (residues 70-270), Brf1 (residues 1-40) and the Bdp1 SANT domain (residues 416-464) and fitted these as well. C34 WH1 and WH2 homology models were based on the NMR structure of the corresponding mouse domains (PDB 2dk8, PDB 2dk5) and generated with Modeller [178] and placed into the density. We then extended the Bdp1 structure by manual modelling in COOT [51]. Towards the C-terminus, we extended the SANT domain with two additional helices that were predicted on a secondary structure

Table 6.2: Data collection and model refinement statistics for the Pol III-TFIIIB complex

	#1 ITC (EMDB-4181) (PDB 6f41)	#2 OC (EMDB-4180) (PDB 6f40)	#4 CC1 EMDB-4182) (PDB 6f42)	#3 CC2 (EMDB-4183) (PDB 6f44)
Data collection and processing				
Magnification	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300
Electron exposure (e–/Å^2)[1]	61.3	60.0/61.8	60.0/61.8	60.0/61.8
Defocus range (μm)	-1 to -3	-0.5 to -4	-0.5 to -4	-0.5 to -4
Pixel size (Å)	1.35	1.35	1.35	1.35
Symmetry imposed	n/a	n/a	n/a	n/a
Initial particle images (no.)	472,519	714,312	714,312	714,312
Final particle images (no.)	29,951	62,751	18,760	34,176
Map resolution (Å)	4.3	3.7	5.5	4.2
FSC threshold	0.143	0.143	0.143	0.143
Map resolution range (Å)	3.9-9.0	3.3-7.5	4.0-9.0	3.7-9.0
Refinement				
Initial model used (PDB code)	5FJ8	5FJ8	5FJ8	5FJ8
Model resolution (Å)	4.3	3.7	5.5	4.2
Map sharpening *B* factor (Å^2)	142	130	100	90
Model composition				
Non-hydrogen atoms	48,614	47,912	45,594	45,573
Protein residues	5,721	5,721	5,515	5,512
Ligands	0	0	0	0
***B* factors (Å^2)**				
Protein	125.93	57.53	254.30	121.87
Ligand	n/a	n/a	n/a	n/a
R.m.s. deviations				
Bond lengths (Å)	0.01	0.01	0.01	0.01
Bond angles (°)	1.03	1.04	1.03	1.03
Validation				
MolProbity score	2.12	2.06	2.18	2.07
Clashscore	8.71	8.23	11.19	8.77
Poor rotamers (%)	1.12	0.96	1.07	0.81
Ramachandran plot				
Favored (%)	87.37	87.17	88.35	88.01
Allowed (%)	12.59	12.72	11.56	11.92
Disallowed (%)	0.04	0.11	0.09	0.07

[1] Structures OC, CC1 and CC2 were determined from two datasets. The electron dose of both sets is given.

level and noticed that a prominent density running in parallel with the Brf1 homology domain II fitted the predicted coiled-coil in Bdp1 perfectly. Towards the N-terminus we manually built the linker and tether regions of Bdp1. Building was aided by secondary structure predictions and bulky side chains that were visible in the EM density. During the preparation of the manuscript, the human TBP-Brf2 (residues 64-407)-Bdp1 (residues 286-407) crystal structure [66] became available, giving additional confidence in assigning the density N-terminal to the Bdp1-SANT domain. Finally, the prominent density running underneath the DNA was assigned to Bdp1 residues 275-320 based on available protein-DNA crosslinks [97, 186] that placed a region of Bdp1 on the opposite side of TBP and was mapped to Bdp1 residues 299-315. Reported photo-crosslinks of Bdp1 K281 to Brf1 and Bdp1 residues 291-295 to the C128 protrusion gave additional confidence [85]. Finally, our cryo-EM density in the 3.7 Å map showed side chain density in this region that allowed me to obtain the sequence register. I also used the 3.4 Å "Pol III PIC joined" map for model building as it showed improved sidechain densities that was helpful especially for building the Bdp1 ER I and Bdp1 tether region.

6.3.5 *Local amplitude scaling (LocScale)*

LocScale maps were calculated as described [88]. In brief, unsharpened and unfiltered cryo-EM maps were scaled against simulated model maps using a rolling window corresponding to seven times the average resolution of the cryo-EM map (19 voxels (25.7 Å) for the OC map, 22 voxels (29.7 Å) for the ITC and CC2 map, and 25 voxels (33.8 Å) for the CC1 map). Model maps were simulated from full-sidechain models refined against cryo-EM maps that were sharpened using the *relion postprocess* program. The dmin parameter of the LocScale program was set to Nyquist frequency (2.7 Å).

6.3.6 *Model refinement and validation*

Models were refined against respective EM maps which were B-factor sharpened with *relion postprocess*. I used a refinement strategy essentially as described previously [82] based on PHENIX [2] libraries. Geometry statistics were calculated with MolProbity [28]. The local resolution was calculated with blocres [80] using a box size of 20 voxels and a FSC cutoff of 0.5. Figures were prepared with UCSF Chimera [157], UCSF ChimeraX and Pymol [39].

6.3.7 *Modelling of the open clamp, closed complex structure*

The open clamp, CC model was obtained by combing the structure of open clamp apo Pol III (PDB 5fja) with the DNA of the human Pol II PIC in its closed state (PDB 5iya), which resembles the DNA in the Pol III OC much more than the yeast Pol II OC. The position of C34 was obtained by superimposing C34 from the Pol III OC model (this work) onto the open clamp structure and aligning the WH3 domains. In this model there are only minor clashes between DNA and the Rpb5 subunit in the jaw which could easily be accommodated by a slightly different curvature of the DNA.

6.3.8 *RNA extension assay*

The RNA extension assay was performed with the DNA oligos and 6-mer RNA described (ITC scaffold). RNA was first radioactively labelled with ^{32}P using T4 polynucleotide kinase and purified over a 15% denaturing urea-polyacrylamide gel. The ITC scaffold was annealed as described above but with the ^{32}P-labelled RNA, and 2 pmol scaffold were pre-incubated with 4 pmol of Pol III for 20 min at 20 °C, followed by incubation with 12 pmol of TFIIIB for 20 min at 20 °C in 20 mM Tris pH 7.5, 200 mM NaCl, 10 mM DTT and 10 mM $MgCl_2$. RNA elongation was initiated by addition of 1 mM ATP, GTP and UTP. After incubation for 10 min at 28 °C, RNA extension was stopped by addition of 0.1%S DS, anf 30 mM EDTA. After phenol extraction and ethanol precipitation, the resulting ^{32}P-labelled RNA products were separated on a denaturing 15 % polyacrylamide-urea gel. RNA bands were detected on an imaging plate (Fujifilm) using a Typhon FLA9500 phosphoimager. The digital image was cropped and contrast was adjusted with the 'levels' tool in Photoshop CS6 v13.0.1.

6.3.9 *Data availability*

Cryo-EM maps of the Pol III ITC, Pol III OC and Pol III CC1 and Pol III CC2 and an intermediate 3.4 Å OC map that aided model building (Pol III PIC joined, see Methods) have been deposited in the Electron Microscopy Data Base (EMDB) under accession codes EMD-4181 (Pol III ITC), EMD-4180 (Pol III OC), EMD-4182 (Pol III CC1) and EMD-4183 (Pol III CC2) and EMD-4184 (Pol III PIC joined). The coordinates of the corresponding atomic models have been deposited in the Protein Data Bank under accession code 6f41 (Pol III ITC), 6f40 (Pol III OC1), 6f42 (Pol III CC1), 6f44 (Pol III CC2). The raw movies were also deposited in the EMPIAR database under accesion numbers EMPIAR-10168 and EMPIAR-10190.

6.4 STRUCTURE DETERMINATION OF THE MAF1-POL III COMPLEX

6.4.1 *Cryo-EM sample preparation*

250 µg of C128-tagged Pol III and a five-fold molar excess of Maf1 i.d (without his-tag) were incubated overnight in binding buffer (40 mM K2 SO4, 15 mM HEPES pH 7.5 and 2.5 mM DTT). The sample was then diluted to 1.8 mL in the same buffer but containing 3 mM of BS3 crosslinker and 2.3 µM Maf1 to prevent dissociation. Sample was crosslinked for 2h at room temperature, quenched with 50 mM Tris pH 7.5, concentrated to ~60 µL and applied to a Superose 6 INCREASE 3.2/300 column equilibrated in EM buffer (150 mM $((NH4)_2)SO_4)$, 15 mM HEPES pH 7.5, 10 mM DTT). A peak fraction eluting at 1.4 mL was diluted to 0.15 mg/mL and used for grid preparation. 2.5 µL of sample were applied to Quanitfoil Cu 2/1 + 2 nm Carbon grids that were glow-discharged in a Pelco EasyGlow instrument. Access sample was blotted away in a Vitrobor Mark IV set to 4 °C and 100 % humidity for 6 seconds with a blot force of 2 and the grid was plunge-frozen in liquid ethane.

6.4.2 *Cryo-EM imaging and pre-processing*

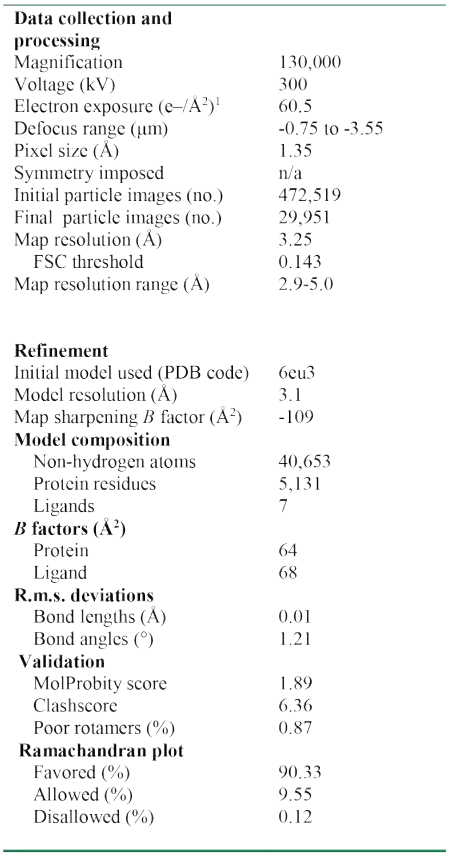
Cryo-EM data was collected on a Titan Krios microscope with a Gatan Quantum energy filter and a K2 Summit direct detector in counting mode at a magnification of 130 000x and a pixel size of 1.041 Å /px. 10520 movies with 40 frames each and an accumulated dose of 60.5 e$^-$ / Å^2 were collected, acquiring 8 shots per hole.
Movies were pre-processed on the fly using Warp 1.05 [194]. The model parameters for motion correction and CTF estimation were set to 5×5×40 and 5×5×1, respectively. Particles were picked with `BoxNet2_20180602` without re-training, using an expected diameter of 250 Å. Particles were inverted, normalized and exported in a 360 px box. Filters in Warp were set to exclude particles from micrographs with lower then 6 Å estimated resolution based in CTF fittimg, with a larger then 1.7 Å movement per frame in the first third of the movie, and with fewer than 50 particles per micrograph. However, very few micrographs did not satisfy these criteria (125 excluded micrographs, ~1% of the collected data), indicating high quality of the sample.

6.4.3 *Maf1-Pol III particle classification*

Due to the large number of particles in the dataset (1 6788 795), the dataset was divided into four batches and each was classified by 3D classification in RELION 3.0 using a 60 Å low-pass filtered model of apo Pol III (PDB 5fj9) as reference. The best class of each batch was retained, yielding a dataset of 728 000 particles. All batches were combined and refined. The resulting map showed density for Maf1 and an adjacent density corresponding to C34 WH2, but at a lower threshold. Using a classification focused on this region this density could be improved (305 000 particles). Finally, a masked classification focussing on the stalk, trimer and clamp seperated an open clamp state, and a closed clamp state, containing 117 442 particles. Density for Maf1 was smeared out in the open clamp state, but clear separation of β-sheets and sidechain density were visible in the closed clamp state. Two rounds of CTF refinement improved the resolution from 3.4 Å to 3.25 Å. However, re-extracting all particles from micrographs that were motion-corrected in RELION 3.0 and employment of the Bayesian Polishing program did not increase the resolution further.
Starting from the 305 000 particle set, I also performed MultiBody refinement implemented in RELION 3, using two masks. The first covered the core of Pol III and the heterodimer, and the second covered the stalk, clamp, heterotrimer and Maf1. The resulting bodies were sharpened using the `relion_postprocess` program, yielding maps at 3.34 Å (core) and 3.74 Å (stalk-trimer-clamp) resolution. This significantly improved side-chain density in the heterotrimer and clamp.

6.4.4 *Model building*

For Model building, PDB 6eu3 (apo-Pol III closed conformation at 3.3 Å resolution) was used as a starting point and combined with the C34 WH2 domain from PDB 6f40. A homology model of yeast Maf1 was generated with PHYRE2 [111] and fitted into the density. The model was manually adjusted in COOT [51].
To improve model coordinates, I used the maps obtained by MultiBody refinement and corrected minor sequence register shifts that were present in the starting model and could be detected due to the higher quality of the maps. Progress of model building



Data collection and processing	
Magnification	130,000
Voltage (kV)	300
Electron exposure (e–/Å^2)^1	60.5
Defocus range (μm)	-0.75 to -3.55
Pixel size (Å)	1.35
Symmetry imposed	n/a
Initial particle images (no.)	472,519
Final particle images (no.)	29,951
Map resolution (Å)	3.25
FSC threshold	0.143
Map resolution range (Å)	2.9-5.0
Refinement	
Initial model used (PDB code)	6eu3
Model resolution (Å)	3.1
Map sharpening *B* factor (Å^2)	-109
Model composition	
Non-hydrogen atoms	40,653
Protein residues	5,131
Ligands	7
***B* factors (Å^2)**	
Protein	64
Ligand	68
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.21
Validation	
MolProbity score	1.89
Clashscore	6.36
Poor rotamers (%)	0.87
Ramachandran plot	
Favored (%)	90.33
Allowed (%)	9.55
Disallowed (%)	0.12

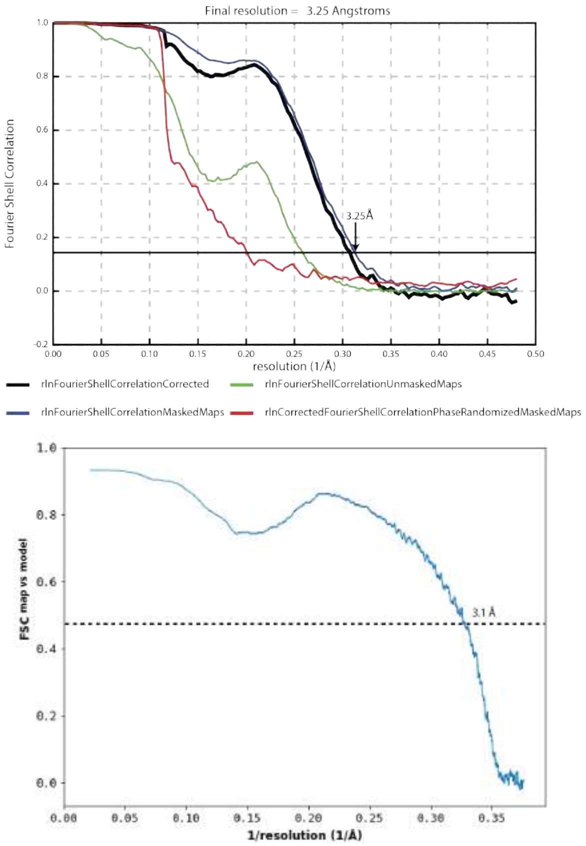


Figure 6.1: Data collection and refinement statistics of the Maf1-Pol III complex

.

was monitored by real-space refinement in PHENIX [2] of partial models, namely the heterodimer, the heterotrimer-clamp region, and Maf1 against respective maps obtained from MultiBody refinement. Finally, the refined models were combined with and refined in another round of real-space refinement against the locally filtered 3.25 Å consensus map obtained using the `relion_postprocess` program.

MolProbity and PHENIX model validation statistics show good stereochemistry and excellent correlation between the map and the model (correlation coefficient of 0.81, and a $FSC_{0.5}$ of the simulated model map with the experimental map of 3.1 Å, see figure 6.1).

6.4.5 *Maf1-C34 pull-down*

180 µg of his-tagged C34(1-156) were incubated with an equimolar amount of Maf1 constructs for 1h, and then diluted tenfold in IP buffer (50 mM NaCl, 20 mM HEPES pH 7.5 and 5 mM DTT). The sample was applied to a small gravity column containing 50 µL of Ni-NTA beads which was then washed with 1 mL of IP buffer. Proteins were eluted in 100 µL of elution buffer (300 mM imidazole, 200 mM NaCl, 50 mM Tris) and analysed by SDS-PAGE.

6.5 TFIIIC

6.5.1 *Reconstitution of the τA -TFIIIB complex*

40 µg τA were incubated at a 1:1 molar ratio with a pre-annealed double-stranded DNA oligo encompassing the upstream TFIIIB-binding region and the A-Box of the yeast gene His_tH(GUG)E2 (Non-template strand: 5'-AGCCCGTAATACAACAGTTCTCCATTGAA-AAGTCGCCATCTTAGTATAGTGGTTAG-'3) for 10 minutes. The sample was diluted with 20 µL of SEC buffer (15 mM HEPES, 5 mM DTT, initally 200 mM KOAc, but later 50 mM KOAc were shown to be required for stochiometric binding of DNA) and incubated with Brf1-TBP at an equimolar ratio for 10 min followed by incbuation with Bdp1 at an equimolar ration for another 20 mins. The sample was then applied to either a Superdex 200 INCREASE 3.2/300 column (when no crosslinking was performed) or to a Superose 6 INCREASE 3.2/300 column when sample was crosslinked.

For the negative stain dataset that gave rise to the interpretable reconstruction, sample was diluted to 0.02 mg/mL (calculated for τA) without prior SEC purification and crosslinked for 30 min on ice with 0.01% glutaraldehyde and quenched woth 50 mM Tris pH 7.5.

For preparation of crosslinked sample for cryo-EM, 200 µg of τA were used. The sample was diluted to 2 mL in SEC buffer and crosslinked with 0.075% τA for 20 min on ice. Crosslinking was quenched by addition of 50 mM Tris pH 7.5 and the sample concentrated to ~60 µL and then fractionated on a Superose 6 3.2/300 column. Sample eluting at ~1.4 mL ws used for grid prepraration at an absorbance of ~0.33 measured on a Nanodrop.

6.5.2 *Negative stain EM structure data collection and processing*

For negative stain grid preparation, 3 µL sample at a concentration of ~0.02 mg/mL were applied to freshly glow-discharged copper grids and incubated for 1 minute. Excess sample was removed with filter papers, the grid was washed twice in 20 µL of buffer, once in 20 µL of 1% uranyl acetate, and then incubated for 1 min with 20 µL 1% uranyl acetate before removing excess tain with filter papers and air-drying the grid for 20 minutes.

For the τA-TFIIIB dataset, 515 micrographs were collected at a magnification of 49 000 corresponding to a magnified mixel size of 2.292 Å / px and a defocus of 0.5 µm on Tecnai Spirit microscope with a Lab6 filament.

Initially, 173 particles were manually picked and subjected to 2D classification (using 10 classes) in RELION 3.0 to generate templates for autopicking. Autopicking of the entire dataset gave 126 645 particle coordinates. Two rounds of 2D classification (with 200 classes each) were performed, and 16 classes comprising 7329 particles were selected used for the RELION Initial model program, using a mask diameter of 200 Å. The obtained model was refined against the same particles to reduce noisyness and overfitting. To obtain more particles, a new round of autopicking was performed, reducing the minimal allowed inter-particle distance to 100 Å, which yielded 356 553 particle corrdinates. Two alternative processing branches were pursued from there (see figure 6.2. For branch A, autopicked particles were first subjected to 3D classification using the refined initial model as reference. The best class (87 978 particles) was cleaned by 2D classification, and 69934 particles were retained. These particles were refined to 18 Å resolution. For branch B, autopicked particles were first subjected to 2D classification and then to 3D classification. The resulting 36 612 particles were refined to 17 Å.

Both maps are similiar, however density for TFIIIB appears more fragmented in the 17 Å map, whereas density for the τ131 TPR array fits the crystal structure better. Density fits displayed in figures are based on the 18 Å map.

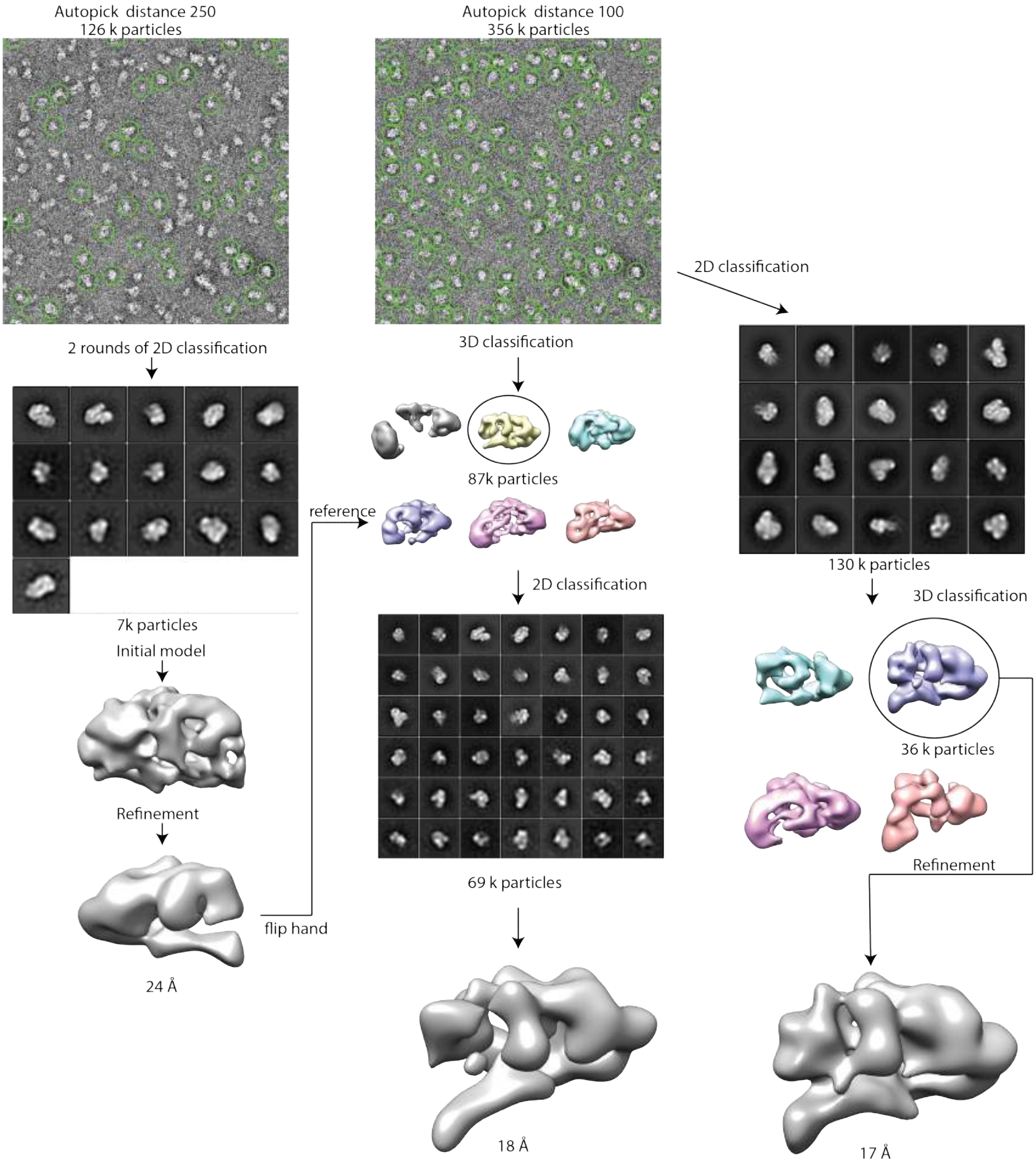

Autopick distance 250
126 k particles
Autopick distance 100
356 k particles
2D classification
2 rounds of 2D classification
3D classification
87k particles
reference
7k particles
Initial model
2D classification
130 k particles
3D classification
36 k particles
Refinement
Refinement
69 k particles
flip hand
24 Å
18 Å
17 Å

Figure 6.2: Data collection and processing for the τA-TFIIIB negative stain structure. Details are given in the text.

6.5.3 *Transcsription assay*

For *in vitro* transcription, 2 pmol of double standed DNA templates were incubated with 10 pmol of transcription factors and 9 pmol Pol III in a 15 µL reaction and incuabted for 10 minutes to allow binding. The reaction was then started by addition of 1 mM each ATP, GTP, CTP and ^{32}P-UTP and allowed to proceede for 20 minutes at 28 °C. The reaction was stopped by addition of 138 µL of deproteination buffer (0.5% SDS, 30 mM EDTA, 0. 6M sodium acetate) and subsequent phenol-chloroform precipitation of nucleic acids. Precipitated nucleic acids were pelleted, washed in 70% ethanol, air-dried and resuspended in 4 µL water and 6 µL formamide. Samples were boiled for 3 min and seperated on a 8% polyacrylamid-urea gel. RNA bands were detected on an imaging plate (Fujifilm) using a Typhon FLA9500 phosphoimager.

Part V

APPENDIX

BIBLIOGRAPHY

1. Abascal-palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553,** 301–306 (2018) (cit. on p. 62).
2. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66,** 213–221 (2010) (cit. on pp. 26, 102, 105).
3. Andrau, J. C., Sentenac, A. & Werner, M. Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. *Journal of Molecular Biology* **288,** 511–520 (1999) (cit. on p. 13).
4. Arimbasseri, A. G. & Maraia, R. J. RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends in Biochemical Sciences* **41,** 546–559 (2016) (cit. on p. 21).
5. Arimbasseri, A. G., Rijal, K. & Maraia, R. J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5,** e27639 (2014) (cit. on pp. 3, 4).
6. Arrebola, R., Manaud, N., Rozenfeld, S., Marsolier, M. C., Lefebvre, O., Carles, C., Thuriaux, P., Conesa, C. & Sentenac, A. Tau91, an essential subunit of yeast transcription factor IIIC, cooperates with tau138 in DNA binding. *Mol Cell Biol* **18,** 1–9 (Jan. 1998) (cit. on p. 10).
7. Bartholomew, B., Kassavetis, G. A., Braun, B. R. & Geiduschek, E. P. The subunit structure of Saccharomyces cerevisiae transcription factor IIIC probed with a novel photocrosslinking reagent. *The EMBO Journal* **9,** 197–2205 (1990) (cit. on pp. 9, 10).
8. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **67,** 271–81 (Apr. 2011) (cit. on p. 57).
9. Belagal, P., Normand, C., Shukla, A., Wang, R., Leger-Silvestre, I., Dez, C., Bhargava, P. & Gadal, O. Decoding the principles underlying the frequency of association with nucleoli for RNA polymerase III-transcribed genes in budding yeast. *Molecular Biology of the Cell* **27,** 3164–3177 (2016) (cit. on p. 19).
10. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nature biotechnology* **22,** 1583–7 (Dec. 2004) (cit. on p. 95).
11. Bernard, G. *et al.* Mutations of POLR3A encoding a catalytic subunit of RNA polymerase Pol III cause a recessive hypomyelinating leukodystrophy. *American Journal of Human Genetics* **89,** 415–23 (2011) (cit. on p. 20).
12. Bhargava, P. Epigenetic regulation of transcription by RNA polymerase III. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1829,** 1015–1025 (2013) (cit. on p. 18).

13. Boguta, M. & Leniewska, E. Novel layers of RNA polymerase III control affecting tRNA gene transcription in eukaryotes. *Open Biology* **7,** 170001 (2017) (cit. on p. 14).
14. Bonhoure, N. *et al.* Loss of the RNA polymerase III repressor MAF1 confers obesity resistance. *Genes and Development* **29,** 934–947 (2015) (cit. on p. 14).
15. Borck, G. *et al.* BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. *Genome research* **25,** 155–166 (2015) (cit. on p. 20).
16. Braun, B. R., Bartholomew, B., Kassavetis, G. A. & Geiduschek, E. Topography of transcription factor complexes on the Saccharomyces cerevisiae 5 S RNA gene. *Journal of Molecular Biology* **228,** 1063–1077 (Dec. 1992) (cit. on p. 10).
17. Bridier-Nahmias, A. *et al.* An RNA polymerase III subunit determines sites of retrotransposon integration. *Science* **348,** 585–588 (2015) (cit. on p. 4).
18. Brow, D. A. & Guthrie, C. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* **334,** 213 (July 1988) (cit. on p. 4).
19. Brueckner, F., Armache, K. J., Cheung, A., Damsma, G. E., Kettenberger, H., Lehmann, E., Sydow, J. & Cramer, P. Structure-function studies of the RNA polymerase II elongation complex. *Acta Crystallographica Section D: Biological Crystallography* **65,** 112–120 (2009) (cit. on p. 5).
20. Brueckner, F., Ortiz, J. & Cramer, P. A movie of the RNA polymerase nucleotide addition cycle. *Current Opinion in Structural Biology* **19,** 294–299 (2009) (cit. on p. 5).
21. Brun, I., Sentenac, A. & Werner, M. Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *The EMBO Journal* **16,** 5730–5741 (1997) (cit. on pp. 6, 13, 50).
22. Burger, K. *et al.* Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *Journal of Biological Chemistry* **285,** 12416–12425 (2010) (cit. on p. 19).
23. Burnol, A. F., Margottin, F., Schultz, P., Marsolier, M. C., Oudet, P. & Burnol, A. F. Basal promoter and enhancer element of yeast U6 snRNA gene. *Journal of Molecular Biology* **233,** 644–658 (1993) (cit. on p. 16).
24. Carter, R. & Drouin, G. 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. *Molecular Biology and Evolution* **27,** 1035–1043 (2010) (cit. on pp. 6, 85).
25. Carter-Timofte, M. E., Hansen, A. F., Christiansen, M., Paludan, S. R. & Mogensen, T. H. Mutations in RNA Polymerase III genes and defective DNA sensing in adults with varicella-zoster virus CNS infection. *Genes and Immunity,* 1–10 (2018) (cit. on p. 21).
26. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: An expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Research* (2016) (cit. on p. 16).
27. Chen, M. & Gartenberg, M. R. Coordination of tRNA transcription with export at nuclear pore complexes in budding yeast. *Genes and Development* **28,** 959–970 (2014) (cit. on p. 19).

28. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* **66,** 12–21 (2010) (cit. on pp. 26, 102).
29. Cheung, S., Ma, L., Chan, P. H., Hu, H. L., Mayor, T., Chen, H. T. & Measday, V. Ty1 integrase interacts with RNA Polymerase III-specific Subcomplexes to Promote Insertion of Ty1 Elements Upstream of Polymerase (Pol) III-transcribed Genes. *Journal of Biological Chemistry* **291,** 6396–6411 (Mar. 2016) (cit. on p. 4).
30. Chiu, Y. H., MacMillan, J. B. & Chen, Z. J. RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway. *Cell* **138,** 576–591 (2009) (cit. on p. 21).
31. Chymkowitch, P. & Enserink, J. M. Regulation of tRNA synthesis by posttranslational modifications of RNA polymerase III subunits. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1861,** 310–319 (2018) (cit. on pp. 15, 27).
32. Ciesla, M. & Boguta, M. Regulation of RNA polymerase III transcription by Maf1 protein. *Acta Biochim. Pol.* **55,** 215–225 (2008) (cit. on p. 14).
33. Ciesla, M., Skowronek, E. & Boguta, M. Function of TFIIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription. *Nucleic acids research* **46,** 9444–9455 (2018) (cit. on pp. 17, 91).
34. Cloutier, T. E., Librizzi, M. D., Mollah, A. K., Brenowitz, M. & Willis, I. M. Kinetic trapping of DNA by transcription factor IIIB. *Proc Natl Acad Sci U S A* **98,** 9581–9586 (2001) (cit. on pp. 9, 46).
35. Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D. & Bogenhagen, D. F. Purified RNA polymerase III accurately and efficiently terminates transcription of 5s RNA genes. *Cell* **34,** 829–835 (1983) (cit. on p. 4).
36. Cramer, P. Structure determination of transient transcription complexes. *Biochemical Society transactions,* 1177–1182 (2016) (cit. on p. 22).
37. Crepaldi, L., Policarpi, C., Coatti, A., Sherlock, W. T., Jongbloets, B. C., Down, T. A. & Riccio, A. Binding of TFIIIC to SINE Elements Controls the Relocation of Activity-Dependent Neuronal Genes to Transcription Factories. *PLoS Genetics* **9** (2013) (cit. on p. 19).
38. Dauwerse, J. G. *et al.* Mutations in genes encoding subunits of RNA polymerases i and III cause Treacher Collins syndrome. *Nature Genetics,* 20–22 (2011) (cit. on p. 20).
39. Delano, W. L. The PyMOL Molecular Graphics System. *Schrodinger* (2002) (cit. on p. 102).
40. Derenzini, M., Montanaro, L. & Treré, D. What the nucleolus says to a tumour pathologist. *Histopathology* **54,** 753–762 (2009) (cit. on p. 19).
41. Desai, N., Lee, J., Upadhya, R., Chu, Y., Moir, R. D. & Willis, I. M. Two steps in Maf1-dependent repression of transcription by RNA polymerase III. *Journal of Biological Chemistry* **280,** 6455–6462 (2005) (cit. on p. 15).
42. Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L. & Ottonello, S. TFIIIC-independent in vitro transcription of yeast tRNA genes. *Journal of Molecular Biology* **299,** 601–613 (2000) (cit. on pp. 16, 71).

43. Dieci, G. & Sentenac, A. Detours and shortcuts to transcription reinitiation. *Trends in Biochemical Sciences* **28,** 202–209 (2003) (cit. on p. 87).

44. Diego, S. & Jolla, L. Absolute mRNA levels and transcriptional initiation rates in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences* **35,** 5208–5212 (1986) (cit. on p. 85).

45. Dienemann, C., Schwalb, B., Schilbach, S. & Cramer, P. Promoter Distortion and Opening in the RNA Polymerase II Cleft. *Molecular Cell* **73,** 1–10 (2018) (cit. on p. 79).

46. Dong, F., Xie, K., Chen, Y., Yang, Y. & Mao, Y. Polycistronic tRNA and CRISPR guide-RNA enables highly efficient multiplexed genome engineering in human cells. *Biochemical and Biophysical Research Communications* **482,** 889–895 (2017) (cit. on p. 5).

47. Donze, D. & Kamakaka, R. T. RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. *The EMBO Journal* **20,** 520–531 (2001) (cit. on p. 18).

48. Doudna, J. A. & Batey, R. T. Structural insights into the signal recognition particle. *Annual review of biochemistry* **73,** 539–557 (2004) (cit. on p. 4).

49. Ducrot, C., Lefebvre, O., Landrieux, E., Guirouilh-Barbat, J., Sentenac, A. & Acker, J. Reconstitution of the yeast RNA polymerase III transcription system with all recombinant factors. *The Journal of Biological Chemistry* **281,** 11685–92 (Apr. 2006) (cit. on pp. 10, 12).

50. Dumay, H., Rubbi, L., Sentenac, A. & Marck, C. Interaction between yeast RNA polymerase III and transcription factor TFIIIC via ABC10α and τ131 subunits. *Journal of Biological Chemistry* **274,** 33462–33468 (1999) (cit. on p. 13).

51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66,** 486–501 (Apr. 2010) (cit. on pp. 26, 100, 104).

52. Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C. & Cramer, P. Structural Basis of RNA Polymerase I Transcription Initiation. *Cell* **169,** 120–131.e22 (2017) (cit. on p. 87).

53. Fan, X., Wang, J., Zhang, X., Yang, Z., Zhang, J.-C., Zhao, L., Peng, H.-L., Lei, J. & Wang, H.-W. Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *bioRxiv* (2018) (cit. on p. 22).

54. Feklistov, A. *et al.* RNA polymerase motions during promoter melting. *Science* **356,** 863–866 (2017) (cit. on pp. 37, 53).

55. Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., Legrand, P., Steuerwald, U. & Müller, C. W. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502,** 644–649 (2013) (cit. on p. 6).

56. Ferri, M. L., Peyroche, G., Siaut, M., Lefebvre, O., Carles, C., Conesa, C. & Sentenac, a. A novel subunit of yeast RNA polymerase III interacts with the TFIIB-related domain of TFIIIB70. *Molecular and cellular biology* **20,** 488–95 (2000) (cit. on p. 13).

57. Filer, D., Thompson, M. A., Takhaveev, V., Dobson, A. J., Kotronaki, I., Green, J. W., Heinemann, M., Tullet, J. M. & Alic, N. RNA polymerase III limits longevity downstream of TORC1. *Nature* **552,** 263–267 (2017) (cit. on p. 14).

58. Fishman, V. *et al.* 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin. *Nucleic acids research,* 1–18 (2018) (cit. on p. 18).

59. Flores, A. *et al.* A protein-protein interaction map of yeast RNA polymerase III. *Proceedings of the National Academy of Sciences* **96,** 7815–7820 (1999) (cit. on pp. 5, 13).

60. French, S. L. *et al.* Visual analysis of the yeast 5S rRNA gene transcriptome: regulation and role of La protein. *Molecular and cellular biology* **28,** 4576–4587 (2008) (cit. on pp. 3, 17, 85).

61. Gaiser, F., Tan, S. & Richmond, T. J. Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. *Journal of Molecular Biology* **302,** 1119–1127 (2000) (cit. on p. 6).

62. Geiduschek, E. P. & Tocchini-Valentini, E. G. Transcription by RNA Polymerase III. *Annual review of biochemistry,* 873–914 (1988) (cit. on p. 16).

63. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. & Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6,** 343–345 (2009) (cit. on p. 95).

64. Glaeser, R. M. & Han, B.-G. Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophysics Reports* **3,** 1–7 (2017) (cit. on pp. 22, 23, 59).

65. Gouge, J., Satia, K., Guthertz, N., Widya, M., Thompson, A. J., Cousin, P., Dergai, O., Hernandez, N. & Vannini, A. Redox Signaling by the RNA Polymerase III TFIIB-Related Factor Brf2. *Cell* **163,** 1375–1387 (2015) (cit. on pp. 9, 20, 44).

66. Gouge, J. *et al.* Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. *Nature Communications* **8,** 1–10 (2017) (cit. on pp. 9, 46, 102).

67. Graczyk, D., Cieśla, M. & Boguta, M. *Regulation of tRNA synthesis by the general transcription factors of RNA polymerase III - TFIIIB and TFIIIC, and by the MAF1 protein* 2018 (cit. on p. 15).

68. Graczyk, D., White, R. J. & Ryan, K. M. Involvement of RNA Polymerase III in Immune Responses. *Molecular and Cellular Biology* **35,** 1848–1859 (2015) (cit. on p. 21).

69. Haeusler, R. A. & Engelke, D. R. Survey and Summary: Spatial organization of transcription by RNA polymerase III. *Nucleic Acids Research* **34,** 4826–4836 (2006) (cit. on p. 19).

70. Hahn, S. Activation and the role of reinitiation in the control of transcription by RNA polymerase II. *Cold Spring Harb Symp Quant Bio* **63,** 181–188 (1998) (cit. on p. 87).

71. Hahn, S. & Roberts, S. The zinc ribbon domains of the general transcription factors TFIIB and Brf: Conserved functional surfaces but different roles in transcription initiation. *Genes and Development* **14,** 719–730 (2000) (cit. on p. 13).

72. Hall, T. A. & Brown, J. W. [4] - The Ribonuclease P Family. *Methods in Enzymology* **341** (ed Nicholson, A. W.) 56–77 (2001) (cit. on p. 4).

73. Han, Y., Yan, C., Hoang, T., Nguyen, D., Jackobel, A. J., Ivanov, I., Knutson, B. A. & He, Y. Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *eLife,* 1–24 (2017) (cit. on p. 87).

74. Hancks, D. C. & Kazazian, H. H. Active human retrotransposons: Variation and disease. *Current Opinion in Genetics and Development* **22,** 191–203 (2012) (cit. on p. 4).

75. Harismendy, O., Gendrel, C.-G., Soularue, P., Gidrol, X., Sentenac, A., Werner, M. & Lefebvre, O. Genome-wide location of yeast RNA polymerase III transcription machinery. *The EMBO Journal* **22,** 4738–47 (2003) (cit. on pp. 17, 71, 91).

76. He, Y., Fang, J., Taatjes, D. J. & Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* **495,** 481–6 (2013) (cit. on p. 9).

77. He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I. & Nogales, E. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533,** 1–21 (2016) (cit. on p. 53).

78. Hein, N., Hannan, K. M., George, A. J., Sanij, E. & Hannan, R. D. The nucleolus: An emerging target for cancer therapy. *Trends in Molecular Medicine* **19,** 643–654 (2013) (cit. on p. 19).

79. Hernandez, N. Small Nuclear RNA Genes: A Model System to Study Fundamental Mechanisms of Transcription. *Journal of Biological Chemistry* **276,** 26733–26736 (2001) (cit. on p. 17).

80. Heymann, J. B. & Belnap, D. M. Bsoft: Image processing and molecular modeling for electron microscopy. *Journal of Structural Biology* **157,** 3–18 (2007) (cit. on p. 102).

81. Hoang, C. & Ferré-D'Amaré, a. R. Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* **107,** 929–39 (Dec. 2001) (cit. on p. 7).

82. Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J. H., Sachse, C. & Muller, C. W. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528,** 231–236 (2015) (cit. on pp. 5–7, 15, 35, 48, 53, 102).

83. Hopper, A. K., Pai, D. A. & Engelke, D. R. Cellular dynamics of tRNAs and their genes. *FEBS Letters* **584,** 310–317 (2010) (cit. on p. 19).

84. Hu, H.-L., Wu, C.-C., Lee, J.-C. & Chen, H.-T. A Region of Bdp1 Necessary for Transcription Initiation That Is Located within the RNA Polymerase III Active Site Cleft. *Molecular and Cellular Biology* **35,** 2831–2840 (2015) (cit. on p. 8).

85. Hu, H.-L., Wu, C.-C., Lee, J.-C. & Chen, H.-T. A region of Bdp1 necessary for transcription initiation that is located within the RNA polymerase III active site cleft. *Molecular and Cellular Biology* **35,** 2831–2840 (2015) (cit. on pp. 6, 13, 41, 46, 102).

86. Huang, Y. & Maraia, R. J. Comparison of the RNA polymerase III transcription machinery in Schizosaccharomyces pombe, Saccharomyces cerevisiae and human. *Nucleic acids research* **29,** 2675–2690 (2001) (cit. on p. 72).

87. Ishiguro, A., Kassavetis, G. A. & Geiduschek, E. P. Essential Roles of Bdp1 , a Subunit of RNA Polymerase III Initiation Factor TFIIIB , in Transcription and tRNA Processing. *Molecular and cellular biology* **22,** 3264–3275 (2002) (cit. on pp. 8, 9).

88. Jakobi, A. J., Wilmanns, M. & Sachse, C. Iterative model-based density improvement yields better atomic structures from cryo-EM maps. *bioRxiv* (2017) (cit. on pp. 24, 41, 56, 102).
89. Jiao, Y. *et al.* Improved maize reference genome with single-molecule technologies. *Nature* (2017) (cit. on p. 4).
90. Joazeiro, C. A. P., Kassavetis, G. A. & Peter Geiduschek, E. Identical Components of Yeast Transcription Factor IIIB Are Required and Sufficient for Transcription of TATA Box-Containing and TATA-Less Genes. *Molecular and Cellular Biology* **14,** 2798–2808 (1994) (cit. on p. 16).
91. Johnson, S. A. S., Dubeau, L. & Johnson, D. L. Enhanced RNA polymerase III-dependent transcription is required for oncogenic transformation. *Journal of Biological Chemistry* **283,** 19184–19191 (2008) (cit. on p. 20).
92. Johnson, S. S., Zhang, C., Fromm, J., Willis, I. M. & Johnson, D. L. Mammalian Maf1 Is a Negative Regulator of Transcription by All Three Nuclear RNA Polymerases. *Molecular Cell* **26,** 367–379 (2007) (cit. on p. 14).
93. Jordán-Pla, A., Gupta, I., De Miguel-Jimnez, L., Steinmetz, L. M., Chávez, S., Pelechano, V. & Prez-Ortín, J. E. Chromatin-dependent regulation of RNA polymerases II and III activity throughout the transcription cycle. *Nucleic Acids Research* **43,** 787–802 (2015) (cit. on p. 85).
94. Jourdain, S., Acker, J., Ducrot, C., Sentenac, A. & Lefebvre, O. The $\tau 95$ subunit of yeast TFIIIC influences upstream and downstream functions of TFIIIC-DNA complexes. *Journal of Biological Chemistry* **278,** 10450–10457 (Mar. 2003) (cit. on p. 11).
95. Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P. & Sigler, P. B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422,** 534–539 (2003) (cit. on pp. 8, 37, 46).
96. Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66,** 125–32 (Feb. 2010) (cit. on p. 57).
97. Kang, J. J., Kang, Y. S. & Stumph, W. E. TFIIIB Subunit Locations on U6 Gene Promoter DNA Mapped by Site-specific Protein-DNA Photo-cross-linking. *FEBS Letters* (2016) (cit. on p. 102).
98. Kassavetis, G. a. & Geiduschek, E. P. Transcription factor TFIIIB and transcription by RNA polymerase III. *Biochemical Society transactions* **34,** 1082–1087 (2006) (cit. on p. 8).
99. Kassavetis, G. a., Letts, G. a. & Geiduschek, E. P. A minimal RNA polymerase III transcription system. *The EMBO Journal* **18(18),** 5042–5051 (1999) (cit. on pp. 7, 46).
100. Kassavetis, G. a., Letts, G. a. & Geiduschek, E. P. The RNA polymerase III transcription initiation factor TFIIIB participates in two steps of promoter opening. *The EMBO Journal* **20,** 2823–2834 (2001) (cit. on pp. 8, 50).
101. Kassavetis, G. A., Bardeleben, C., Kumar, A., Ramirez, E. & Geiduschek, E. P. Domains of the Brf component of RNA polymerase III transcription factor IIIB (TFIIIB): functions in assembly of TFIIIB-DNA complexes and recruitment of RNA polymerase to the promoter. *Molecular and cellular biology* **17,** 5299–306 (1997) (cit. on p. 46).

102. Kassavetis, G. A., Blanco, J. A., Johnson, T. E. & Geiduschek, E. P. Formation of Open and Elongating Transcription by RNA Polymerase III Complexes. *Journal of Molecular Biology*, 47–58 (1992) (cit. on pp. 12, 13).

103. Kassavetis, G. A., Braun, B. R., Nguyen, L. H. & Peter Geiduschek, E. S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* **60,** 235–245 (1990) (cit. on pp. 13, 46).

104. Kassavetis, G. A., Han, S., Naji, S. & Peter Geiduschek, E. The role of transcription initiation factor IIIB subunits in promoter opening probed by photochemical cross-linking. *Journal of Biological Chemistry* **278,** 17912–17917 (2003) (cit. on p. 13).

105. Kassavetis, G. A., Joazeiro, C. A., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S. & Blanco, J. A. The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. *Cell* **71,** 1055–1064 (1992) (cit. on p. 12).

106. Kassavetis, G. A., Kumar, A., Ramirez, E. & Geiduschek, E. P. Functional and structural organization of Brf, the TFIIB-related component of the RNA polymerase III transcription initiation complex. *Molecular and cellular biology* **18,** 5587–5599 (1998) (cit. on p. 13).

107. Kassavetis, G. a., Prakash, P. & Shim, E. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *Journal of Biological Chemistry* **285,** 2695–2706 (2010) (cit. on pp. 6, 41).

108. Kassavetis, G. A., Soragni, E., Driscoll, R. & Geiduschek, E. P. Reconfiguring the connectivity of a multiprotein complex: fusions of yeast TATA-binding protein with Brf1, and the function of transcription factor IIIB. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 15406–15411 (2005) (cit. on pp. 8, 31, 99).

109. Kedinger, C., Gniazdowski, M., Mandel, J. L., Gissinger, F. & Chambon, P. α-Amanitin: A specific inhibitor of one of two DNA-dependent RNA polymerase activities from calf thymus. *Biochemical and Biophysical Research Communications* **38,** 165–171 (1970) (cit. on p. 17).

110. Keener, J., Josaitis, C. A., Dodd, J. A. & Nomura, M. Reconstitution of yeast RNA polymerase I transcription in vitro from purified components: TATA-binding protein is not required for basal transcription. *Journal of Biological Chemistry* (1998) (cit. on p. 7).

111. Kelly, L., Mezulis, S., Yates, C., Wass, M. & Sternberg, M. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nature Protocols* **10,** 845–858 (2015) (cit. on pp. 100, 104).

112. Khatter, H., Vorländer, M. K. & Müller, C. W. RNA polymerase I and III: similar yet unique. *Current Opinion in Structural Biology* **47,** 88–94 (2017) (cit. on p. 7).

113. Khoo, B., Brophy, B. & Jackson, S. P. Conserved functional domains of the RNA polymerase III general transcription factor BRF. *Genes and Development* **8,** 2879–2890 (1994) (cit. on p. 8).

114. Khoo, S. K., Wu, C. C., Lin, Y. C. & Chen, H. T. The TFIIE-related Rpc82 subunit of RNA polymerase III interacts with the TFIIB-related transcription factor Brf1 and the polymerase cleft for transcription initiation. *Nucleic Acids Research* **46,** 1157–1166 (2018) (cit. on p. 13).

115. Khoo, S.-K., Wu, C.-C., Lin, Y.-C., Lee, J.-C. & Chen, H.-T. Mapping the protein interaction network for TFIIB-related factor Brf1 in the RNA polymerase III preinitiation complex. *Molecular and cellular biology* **34,** 551–9 (2014) (cit. on p. 13).

116. Khoo, S.-K., Wu, C.-C., Lin, Y.-C., Lee, J.-C. & Chen, H.-T. Mapping the protein interaction network for TFIIB-related factor Brf1 in the RNA polymerase III preinitiation complex. *Molecular and cellular biology* **34,** 551–9 (2014) (cit. on p. 44).

117. Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUS in RELION-2. *eLife* **5,** 1–21 (2016) (cit. on p. 99).

118. Kostrewa, D., Zeller, M. E., Armache, K.-j., Seizl, M., Leike, K., Thomm, M. & Cramer, P. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462,** 323–330 (2009) (cit. on p. 8).

119. Kramerov, D. A. & Vassetzky, N. S. *SINEs* 2011 (cit. on p. 4).

120. Kumar, A., Kassavetis, G. A., Geiduschek, E. P., Hambalko, M. & Brent, C. J. Functional dissection of the B'' component of RNA polymerase III transcription factor IIIB: a scaffolding protein with multiple roles in assembly and initiation of transcription. *Molecular and cellular biology* **17,** 1868–1880 (1997) (cit. on pp. 8, 46).

121. Kwapisz, M., Smagowicz, W. J., Oficjalska, D., Hatin, I., Rousset, J. P., Zoładek, T. & Boguta, M. Up-regulation of tRNA biosynthesis affects translational readthrough in maf1-Δ mutant of Saccharomytes cerevisiae. *Current Genetics* **42,** 147–152 (2002) (cit. on p. 14).

122. Lam, Y. W. & Trinkle-, L. The Nucleolus. *Cold Spring Harbor perspectives in biology,* 1335–1337 (2011) (cit. on p. 19).

123. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* (2001) (cit. on p. 4).

124. Lassar, A. B., Martin, P. L. & Roeder, R. G. Transcription of class III genes: Formation of preinitiation complexes. *Science* **222,** 740–748 (1983) (cit. on p. 87).

125. Layat, E., Probst, A. V. & Tourmente, S. Structure, function and regulation of Transcription Factor IIIA: From Xenopus to Arabidopsis. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1829,** 274–282 (2013) (cit. on p. 16).

126. Lefebvre, O., Ruth, J. & Sentenac, A. A mutation in the largest subunit of yeast TFIIIC affects tRNA and 5 S RNA synthesis. Identification of two classes of suppressors. *Journal of Biological Chemistry* **269,** 23374–23381 (1994) (cit. on p. 10).

127. Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research* **46,** 493–496 (2018) (cit. on p. 8).

128. Liao, Y., Moir, R. D. & Willis, I. M. Interactions of Brf1 peptides with the tetratricopeptide repeat-containing subunit of TFIIIC inhibit and promote preinitiation complex assembly. *Molecular and cellular biology* **26,** 5946–56 (2006) (cit. on p. 12).

129. Liao, Y., Willis, I. M. & Moir, R. D. The Brf1 and Bdp1 Subunits of Transcription Factor TFIIIB Bind to Overlapping Sites in the Tetratricopeptide Repeats of Tfc4. *Journal of Biological Chemistry* **278,** 44467–44474 (Nov. 2003) (cit. on p. 12).

130. Lockwood, W. W. *et al.* Integrative genomic analyses identify BRF2 as a novel lineage-specific oncogene in lung squamous cell carcinoma. *PLoS Medicine* **7** (2010) (cit. on p. 20).

131. Loewith, R. & Hall, M. N. Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* **189,** 1177–1201 (2011) (cit. on p. 14).

132. Louder, R. K., He, Y., López-Blanco, J. R., Fang, J., Chacón, P. & Nogales, E. Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature* **531,** 604–609 (2016) (cit. on p. 90).

133. Lu, P. *et al.* Three-dimensional structure of human γ-secretase. *Nature* **512,** 166–170 (June 2014) (cit. on p. 20).

134. Male, G., von Appen, A., Glatt, S., Taylor, N. M. I., Cristovao, M., Groetsch, H., Beck, M. & Müller, C. W. Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly. *Nature Communications* **6,** 7387 (2015) (cit. on pp. 9–12, 96).

135. Manaud, N., Arrebola, R., Buffin-Meyer, B., Lefebvre, O., Voss, H., Riva, M., Conesa, C. & Sentenac, A. A Chimeric Subunit of Yeast Transcription Factor IIIC Forms a Subcomplex with τ95. *Molecular and Cellular Biology* **18,** 191–200 (1998) (cit. on p. 10).

136. Marzluff, W. F. & Huang, R. C. C. Chromatin Directed Transcription of 5S and tRNA Genes. *Proceedings of the National Academy of Sciences of the United States of America* **72,** 1082–1086 (1975) (cit. on pp. 5, 17).

137. Marzouki, N., Camier, S., Ruet, A., Moenne, A. & Sentenac, A. Selective proteolysis defines two DNA binding domains in yeast transcription factor τ. *Nature* **323,** 176–178 (1986) (cit. on p. 9).

138. Mattison, J. A. *et al.* Caloric restriction improves health and survival of rhesus monkeys. *Nature Communications* **8,** 1–12 (2017) (cit. on p. 14).

139. Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V. & Adkins, J. N. Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Science* **23,** 747–759 (2014) (cit. on p. 59).

140. Misa, J., Schwartz, C. & Wheeldon, I. Design of Hybrid RNA Polymerase III Promoters for Efficient CRISPR-Cas9 Function. *BIO-PROTOCOL* **8** (2018) (cit. on p. 5).

141. Moir, R. D., Puglia, K. V. & Willis, I. M. Interactions between the tetratricopeptide repeat-containing transcription factor TFIIIC131 and its ligand, TFIIIB70. Evidence for a conformational change in the complex. *The Journal of Biological Chemistry* **275,** 26591–8 (Aug. 2000) (cit. on pp. 9, 12).

142. Moir, R. D., Lee, J. & Willis, I. M. Recovery of RNA polymerase III transcription from the glycerol-repressed state. *Journal of Biological Chemistry* **287,** 30833–30841 (2012) (cit. on pp. 14, 15, 55).

143. Moir, R. D., Puglia, K. V. & Willis, I. M. Autoinhibition of TFIIIB70 binding by the tetratricopeptide repeat-containing subunit of TFIIIC. *Journal of Biological Chemistry* **277,** 694–701 (Jan. 2002) (cit. on p. 12).

144. Moir, R. D., Sethy-Coraci, I., Puglia, K., Librizzi, M. D. & Willis, I. M. A tetratricopeptide repeat mutation in yeast transcription factor IIIC131 (TFIIIC131) facilitates recruitment of TFIIB-related factor TFIIIB70. *Mol Cell Biol* **17,** 7119–25. ST –A tetratricopeptide repeat mutation (1997) (cit. on p. 12).

145. Moir, R. D. & Willis, I. M. Regulation of pol III transcription by nutrient and stress signaling pathways. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1829,** 361–375 (2013) (cit. on pp. 14, 85).

146. Moqtaderi, Z. & Struhl, K. Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces cerevisiae reveals loci with incomplete transcription complexes. *Molecular and cellular biology* **24,** 4118–27 (2004) (cit. on pp. 17, 18, 71).

147. Moqtaderi, Z., Wang, J., Raha, D., White, R. J., Snyder, M., Weng, Z. & Struhl, K. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nature Structural and Molecular Biology* **17,** 635–640 (2010) (cit. on pp. 16–18).

148. Moreno-Morcillo, M., Taylor, N. M., Gruene, T., Legrand, P., Rashid, U. J., Ruiz, F. M., Steuerwald, U., Müller, C. W. & Fernández-Tornero, C. Solving the RNA polymerase I structural puzzle. *Acta Crystallographica Section D: Biological Crystallography* **70,** 2570–2582 (2014) (cit. on p. 95).

149. Mylona, A., Fernández-Tornero, C., Legrand, P., Haupt, M., Sentenac, A., Acker, J. & Müller, C. W. Structure of the $\tau60/\Delta\tau91$ Subcomplex of Yeast Transcription Factor IIIC: Insights into Preinitiation Complex Assembly. *Molecular Cell* **24,** 221–232 (2006) (cit. on pp. 10, 12).

150. Nogales, E., Louder, R. K. & He, Y. Cryo-EM in the study of challenging systems: the human transcription pre-initiation complex. *Current opinion in structural biology,* 120–127 (2016) (cit. on p. 22).

151. Oficjalska-Pham, D., Harismendy, O., Smagowicz, W. J., Gonzalez de Peredo, A., Boguta, M., Sentenac, A. & Lefebvre, O. General Repression of RNA Polymerase III Transcription Is Triggered by Protein Phosphatase Type 2A-Mediated Dephosphorylation of Maf1. *Molecular Cell* **22,** 623–632 (2006) (cit. on pp. 14, 15).

152. Ogunjimi, B. *et al.* Inborn errors in RNA polymerase III underlie severe varicella zoster virus infections. *Journal of Clinical Investigation* **127,** 3543–3556 (2017) (cit. on p. 21).

153. Oler, A. J. *et al.* Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nature Structural and Molecular Biology* **17,** 620–628 (2010) (cit. on pp. 17, 18).

154. Papantonis, A. & Cook, P. R. Transcription factories: Genome organization and gene regulation. *Chemical Reviews* **113,** 8683–8705 (2013) (cit. on p. 18).

155. Park, J. L., Lee, Y. S., Kunkeaw, N., Kim, S. Y., Kim, I. H. & Lee, Y. S. Epigenetic regulation of noncoding RNA transcription by mammalian RNA polymerase III. *Epigenomics* **9,** 171–187 (2017) (cit. on p. 4).

156. Pelechano, V., Chavez, S. & Perez-Ortin, J. E. A Complete Set of Nascent Transcription Rates for Yeast Genes. *PLoS ONE* **5,** e15442 (2010) (cit. on pp. 3, 85).

157. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25,** 1605–1612 (2004) (cit. on p. 102).

158. Pianese, G. Beitrag zur Histologie und Aetiologie der Carcinoma. Histologische und experimentelle Untersuchungen. *Beitr. Pathol. Anat. Allg Pathol.* **142,** 1–193 (1896) (cit. on p. 19).

159. Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J. & Cramer, P. Transcription initiation complex structures elucidate DNA opening. *Nature* **533,** 353–8 (2016) (cit. on p. 53).

160. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14,** 290–296 (2017) (cit. on p. 24).

161. Raab, J. R., Chiu, J., Zhu, J., Katzman, S., Kurukuti, S., Wade, P. A., Haussler, D. & Kamakaka, R. T. Human tRNA genes function as chromatin insulators. *The EMBO Journal* **31,** 330–350 (2012) (cit. on p. 18).

162. Rani, P. G., Ranish, J. A., Hahn, S., Hughes, H. & Al, R. E. T. RNA Polymerase II (Pol II) -TFIIF and Pol II-Mediator Complexes : the Major Stable Pol II Complexes and Their Activity in Transcription Initiation and Reinitiation. *Molecular and Cellular Biology* **24,** 1709–1720 (2004) (cit. on p. 87).

163. Reina, J. H., Azzouz, T. N. & Hernandez, N. Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS ONE* **1** (2006) (cit. on p. 15).

164. Rijal, K. & Maraia, R. J. RNA polymerase III mutants in TFIIFa-like C37 that cause terminator readthrough with no decrease in transcription output. *Nucleic Acids Research* **41,** 139–155 (2013) (cit. on pp. 6, 41).

165. Riva, M., Schultz, P. & Carles, C. The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is, 3857–3871 (1998) (cit. on p. 6).

166. Roberts, D. N., Stewart, A. J., Huff, J. T. & Cairns, B. R. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 14695–700 (2003) (cit. on pp. 17, 18).

167. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224,** 488–490 (1969) (cit. on p. 5).

168. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology* **333,** 721–745 (2003) (cit. on p. 24).

169. Ross, A. P. & Zarbalis, K. S. The emerging roles of ribosome biogenesis in craniofacial development. *Frontiers in Physiology* **5** (2014) (cit. on p. 20).

170. Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5 end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. *Cell* **54,** 795–804 (1988) (cit. on p. 46).

171. Roy, K. & Chanfreau, G. F. A global function for transcription factors in assisting RNA polymerase II termination. *Transcription* **9,** 41–46 (2018) (cit. on p. 19).

172. Roy, K., Gabunilas, J., Gillespie, A., Ngo, D. & Chanfreau, G. F. Common genomic elements promote transcriptional and DNA replication roadblocks. *Genome Research* **26,** 1363–1375 (2016) (cit. on pp. 19, 46).

173. Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D. & Cramer, P. Evolution of two modes of intrinsic RNA polymerase transcript cleavage. *Journal of Biological Chemistry* (2011) (cit. on p. 6).

174. Sadian, Y. *et al.* Structural insights into transcription initiation by yeast RNA polymerase I. *The EMBO Journal* **36,** 2698–2709 (2017) (cit. on p. 87).

175. Saïda, F. Structural characterization of the interaction between TFIIIB components Bdp1 and Brf1. *Biochemistry* **47,** 13197–13206 (2008) (cit. on p. 8).

176. Sainsbury, S., Niesser, J. & Cramer, P. Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* **493,** 2–6 (2012) (cit. on p. 8).

177. Salem, M., Mauguen, Y. & Prangé, T. Revisiting glutaraldehyde cross-linking: The case of the Arg-Lys intermolecular doublet. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **66,** 225–228 (2010) (cit. on p. 59).

178. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* **234,** 779–815 (1993) (cit. on p. 100).

179. Scheres, S. H. W. Semi-automated selection of cryo-EM particles in RELION-1.3. *Journal of structural biology* **189,** 114–122 (Dec. 2014) (cit. on p. 23).

180. Schmitz, J. in *Repetitive DNA* 92–107 (2012) (cit. on p. 4).

181. Schneider, D. A. RNA polymerase I activity is regulated at multiple steps in the transcription cycle: Recent insights into factors that influence transcription elongation. *Gene* **493,** 176–184 (2012) (cit. on p. 15).

182. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes and Development* **16,** 2593–2620 (2002) (cit. on p. 16).

183. Schultz, P., Marzouki, N., Marck, C., Ruet, A., Oudet, P. & Sentenac, A. The two DNA-binding domains of yeast transcription factor as observed by scanning transmission electron microscopy. **8,** 3815–3824 (1989) (cit. on p. 9).

184. Schwartz, C. M., Hussain, M. S., Blenner, M. & Wheeldon, I. Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica. *ACS Synthetic Biology* **5,** 356–359 (2016) (cit. on p. 5).

185. Sethy-Coraci, I., Moir, R. D., López-de-León, A. & Willis, I. M. A differential response of wild type and mutant promoters to TFIIIB70overexpression in vivo and in vitro. *Nucleic Acids Research* **26,** 2344–2352 (1998) (cit. on p. 12).

186. Shah, S. M. A., Kumar, A., Geiduschek, E. P. & Kassavetis, G. A. Alignment of the B" Subunit of RNA Polymerase III Transcription Factor IIIB in Its Promoter Complex. *Journal of Biological Chemistry* **274,** 28736–28744 (1999) (cit. on pp. 8, 9, 102).

187. Si, D. & He, J. Tracing Beta Strands Using StrandTwister from Cryo-EM Density Maps at Medium Resolutions. *Structure* **22,** 1665–1676 (2014) (cit. on p. 26).

188. Soprano, A. S. *et al.* Crystal Structure and Regulation of the Citrus Pol III Repressor MAF1 by Auxin and Phosphorylation. *Structure* **25,** 1360–1370.e4 (2017) (cit. on p. 15).

189. Soragni, E. & Kassavetis, G. A. Absolute gene occupancies by RNA polymerase III, TFIIIB, and TFIIIC in Saccharomyces cerevisiae. *Journal of Biological Chemistry* **283,** 26568–26576 (2008) (cit. on p. 17).

190. Tafur, L., Sadian, Y., Wetzel, R., Weis, F. & Muller, C. W. Structural rearrangement of TFIIS- and TFIIF/TFIIE-like subunits in RNA polymerase I transcription complexes. *bioRxiv* (2018) (cit. on p. 6).

191. Taylor, K. A. & Glaeser, R. M. Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. *Journal of structural biology* **163,** 214–223 (2008) (cit. on p. 22).

192. Taylor, N. M., Baudin, F., Von Scheven, G. & Müller, C. W. RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74. *Nucleic Acids Research* **41,** 9183–9196 (2013) (cit. on p. 10).

193. Taylor, N. M. *et al.* Structural and functional characterization of a phosphatase domain within yeast general transcription factor IIIC. *Journal of Biological Chemistry* **288,** 15110–15120 (2013) (cit. on p. 10).

194. Tegunov, D. & Cramer, P. *Real-time cryo-EM data pre-processing with Warp* 2018 (cit. on pp. 23, 24, 26, 104).

195. Teichmann, M., Wang, Z. & Roeder, R. G. A stable complex of a novel transcription factor IIB- related factor, human TFIIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 14200–5 (Dec. 2000) (cit. on p. 8).

196. Tse, W. K. F. Treacher Collins syndrome: New insights from animal models. *International Journal of Biochemistry and Cell Biology,* 44–47 (2016) (cit. on p. 20).

197. Turowski, T. W., Leśniewska, E., Delan-Forino, C., Sayou, C., Boguta, M. & Tollervey, D. Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome Research* **26,** 933–944 (2016) (cit. on p. 14).

198. Upadhya, R., Lee, J. H. & Willis, I. M. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Molecular Cell* **10,** 1489–1494 (2002) (cit. on p. 14).

199. Van Bortle, K. & Corces, V. G. tDNA insulators and the emerging role of TFIIIC in genome organization. *Transcription* **3,** 277–284 (2012) (cit. on p. 18).

200. Van Den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *Journal of Biochemical and Biophysical Methods* **67,** 67–74 (2006) (cit. on p. 95).

201. Vannini, A. & Cramer, P. Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *Mol Cell* **45,** 439–446 (2012) (cit. on pp. 6, 56, 62, 65, 85, 87).

202. Vannini, A., Ringel, R., Kusser, A. G., Berninghausen, O., Kassavetis, G. A. & Cramer, P. Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell* **143,** 59–70 (2010) (cit. on p. 15).

203. Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553,** 295–300 (2018) (cit. on pp. 37–40, 42, 47, 49–52, 54, 86, 88).

204. Walker-Kopp, N., Jackobel, A. J., Pannafino, G. N., Morocho, P. A., Xu, X. & Knutson, B. A. Treacher Collins syndrome mutations in Saccharomyces cerevisiae destabilize RNA polymerase I and III complex integrity. *Human Molecular Genetics* **26,** 4290–4300 (2017) (cit. on p. 20).

205. Wei, Y.-Y. & Chen, H.-T. Functions of the TFIIE-related tandem winged-helix domain of Rpc34 in RNA polymerase III initiation and elongation. *Molecular and Cellular Biology,* MCB.00105–17 (2017) (cit. on pp. 13, 50).

206. Werner, M., Chaussivert, N., Willis, I. M. & Sentenac, A. Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. *Journal of Biological Chemistry* **268,** 20721–20724 (1993) (cit. on p. 13).

207. Westover, K. D. Structural Basis of Transcription: Separation of RNA from DNA by RNA Polymerase II. *Science* **303,** 1014–1016 (2004) (cit. on p. 5).

208. White, R. J. RNA polymerase III transcription and cancer. *Oncogene* **23,** 3208–16 (2004) (cit. on p. 14).

209. White, R. J. RNA polymerases I and III, growth control and cancer. *Nature reviews. Molecular cell biology* **6,** 69–78 (2005) (cit. on p. 19).

210. White, R. J. Transcription by RNA polymerase III: more complex than we thought. *Nature reviews. Genetics* **12,** 459–463 (2011) (cit. on p. 17).

211. Willis, I. M. Maf1 phenotypes and cell physiology. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1861,** 330–337 (2018) (cit. on p. 14).

212. Willis, I. M., Moir, R. D. & Hernandez, N. Metabolic programming a lean phenotype by deregulation of RNA polymerase III. *Proceedings of the National Academy of Sciences* **115,** 12182–12187 (2018) (cit. on p. 14).

213. Wu, C. C., Lin, Y. C. & Chen, H. T. The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIIC, Bdp1, and the RNA polymerase III active center. *Mol Cell Biol* **31,** 2715–2728 (2011) (cit. on p. 13).

214. Wu, S. M. & Hochedlinger, K. Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nature cell biology* **13,** 497–505 (May 2011) (cit. on pp. 6, 41, 46).

215. Yi, Y. *et al.* Exploring a common mechanism of alcohol-induced deregulation of RNA Pol III genes in liver and breast cells. *Gene* **626,** 309–318 (2017) (cit. on p. 19).

216. Yudkovsky, N., Ranish, J. A. & Hahn, S. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408,** 225–229 (2000) (cit. on p. 87).

217. Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of Structural Biology* **193,** 1–12 (2016) (cit. on p. 99).

218. Zhang, L., Silva, D. A., Pardo-Avila, F., Wang, D. & Huang, X. Structural Model of RNA Polymerase II Elongation Complex with Complete Transcription Bubble Reveals NTP Entry Routes. *PLoS Computational Biology* (2015) (cit. on p. 5).

219. Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y. & Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14,** 331–332 (2017) (cit. on pp. 23, 99).

220. Zhong, Q., Shi, G., Zhang, Q., Lu, L., Levy, D. & Zhong, S. Tamoxifen represses alcohol-induced transcription of RNA polymerase III-dependent genes in breast cancer cells. *Oncotarget* **5,** 12410–12417 (2014) (cit. on p. 19).

221. Zhong, Q. *et al.* The significance of Brf1 overexpression in human hepatocellular carcinoma. *Oncotarget* **7,** 6243–6254 (2016) (cit. on p. 19).

222. Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. & Scheres, S. H. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7,** e42166 (Nov. 2018) (cit. on p. 23).