Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Wolfram Heinrich Seifert Davila

born in Cajamarca, Peru

Oral examination: 25th of March 2024

Structural and functional studies of human and yeast transcription factor IIIC (TFIIIC)

Referees:

Prof. Dr. Michael Knop

Dr. Sebastian Eustermann

Summary

RNA polymerase III synthesizes various small non-coding RNAs, including 5S rRNA, tRNA, VAI, and U6 small nuclear RNA, which require different promoter types and auxiliary transcription factors. Transcription of tRNA genes occurs from type 2 promoters and is mediated by transcription factors (TF) IIIB and TFIIIC. Recognition of intragenic A- and B-box motifs in tRNA genes by TFIIIC subcomplexes τA and τB is the first critical step of tRNA synthesis but is mechanistically poorly understood. Despite existing structural information on some individual TFIIIC subunits and τA subcomplex, comprehensive structural data on the complete TFIIIC complex and its interaction with promoter regions was lacking. My work bridges this gap by presenting, for the first time, high-resolution cryo-EM structures of both human and yeast TFIIIC complexes.

In this study, I elucidated the structures of human TFIIIC with and without DNA, achieving high-resolution maps ranging from 3.2 to 3.5 Å. These findings unveil the subunit composition of τA and τB subcomplexes in high resolution. A significant revelation was the identification of TFIIIC220's critical role in both subcomplexes and the localization of the flexible linker, which facilitates TFIIIC's binding to its distinct A- and B-box promoters. This was determined through an in-depth analysis of cryo-EM data, including measurements of the distances between τA subcomplex and τB-dimer particles. Subsequently, I proposed a mechanistic model of TFIIIC's interaction with tRNA genes, suggesting that τB, upon recognizing the B-box, anchors the complex to the DNA, thereby enabling the τA subcomplex to engage its promoter through a fly-casting mechanism.

Furthermore, I solved the structures of yeast TFIIIC fully bound to DNA at resolutions of 2.67 Å (τB –DNA complex) and 3.73 Å (τA-DNA subcomplex). These structures not only corroborate a similar promoter recognition mechanism as observed in humans but also provide novel insights. The τB-DNA complex sheds light on promoter recognition, and the function of the flexible linker, akin to TFIIIC220 in humans, is observed in τ138. The τA-DNA complex, for the first time, provides insights into this highly dynamic interaction. This interaction was further validated through single-molecule fluorescent microscopy experiments.

Overall, my work provides unprecedented structural insights into the TFIIIC complex, significantly advancing our understanding of its interaction with DNA and its role in the transcription of tRNA genes.

ZUSAMMENFASSUNG

RNA-Polymerase III synthetisiert verschiedene kleine nicht-kodierende RNAs, einschließlich 5S rRNA, tRNA, VAI und U6 kleine Kern-RNA, die unterschiedliche Promotortypen und Hilfstranskriptionsfaktoren erfordern. Die Transkription von tRNA-Genen erfolgt von Typ-2-Promotoren aus und wird durch die Transkriptionsfaktoren (TF) IIIB und TFIIIC vermittelt. Die Erkennung von intragenischen A- und B-Box-Motiven in tRNA-Genen durch TFIIIC-Subkomplexe τA und τB ist der erste kritische Schritt der tRNA-Synthese, ist aber mechanistisch schlecht verstanden. Trotz vorhandener struktureller Informationen über einige einzelne TFIIIC-Untereinheiten und den τA-Subkomplex fehlten umfassende strukturelle Daten über den vollständigen TFIIIC-Komplex und seine Interaktion mit Promotorregionen. Meine Arbeit schließt diese Lücke, indem sie erstmals hochauflösende Kryo-EM-Strukturen der TFIIIC-Komplexe von Mensch und Hefe präsentiert.

In dieser Studie habe ich die Strukturen des menschlichen TFIIIC mit und ohne gebundene DNA aufgeklärt und hochauflösende Karten im Bereich von 3.2 bis 3.5 Å erreicht. Diese Ergebnisse enthüllen die Untereinheitenzusammensetzung der τA- und τB-Subkomplexe in hoher Auflösung. Eine bedeutende Entdeckung war die Identifizierung der kritischen Rolle von TFIIIC220 in beiden Subkomplexen und die Lokalisierung des flexiblen Linkers, der die Bindung von TFIIIC an seine verschiedenen A- und B-Box-Promotoren erleichtert. Dies wurde durch eine eingehende Analyse der Kryo-EM-Daten bestimmt, einschließlich der Messung der Abstände zwischen τA-Subkomplex und τB-Dimer-Partikeln. Anschließend schlug ich ein mechanistisches Modell der Interaktion von TFIIIC mit tRNA-Genen vor, das darauf hindeutet, dass τB nach Erkennung der B-Box den Komplex an die DNA verankert, wodurch der τA-Subkomplex seinen Promotor durch einen Fly-Casting-Mechanismus binden kann.

Darüber hinaus habe ich die Strukturen des Hefe-TFIIIC vollständig an DNA gebunden bei Auflösungen von 2.67 Å (τB –DNA-Komplex) und 3.73 Å (τA-DNA-Subkomplex) aufgelöst. Diese Strukturen bestätigen nicht nur einen ähnlichen Promotorerkennungsmechanismus wie beim Menschen, sondern liefern auch neue Einblicke. Der τB-DNA-Komplex liefert weitere Erkenntnisse über die Promotorerkennung, und die Funktion des flexiblen Linkers, ähnlich dem TFIIIC220 beim Menschen, wird in τ138 beobachtet. Der τA-DNA-Komplex bietet erstmals Einblicke in diese sehr dynamische Interaktion. Diese Interaktion wurde durch Experimente mit Einzelmolekül-Fluoreszenzmikroskopie weiter validiert.

Insgesamt liefert meine Arbeit beispiellose strukturelle Einblicke in den TFIIIC-Komplex und erweitert erheblich unser Verständnis seiner Interaktion mit DNA und seiner Rolle bei der Transkription von tRNA-Genen.

Peer-reviewed articles

W. Seifert-Davila*, M. Girbig*, L. Hauptmann, T. Hoffmann, S. Eustermann, and C. W. Müller, "Structural insights into human TFIIIC promoter recognition," *Sci. Adv.*, vol. 9, no. 27, pp. 1–11, 2023, doi: 10.1126/sciadv.adh2019

*Equal contribution

Manuscripts in preparation

Seifert-Davila. W, Chaban. A, Girbig. M, Hoffmann. T, Eustermann. S, Duss. O, Müller. C.W. "Cryo-EM structural insights into yeast TFIIIC-DNA interactions enriched by single-molecule fluorescence microscopy analysis"

Acknowledgments

Finally, my partner in crime, Oliwia Koczy, deserves immense gratitude for her unwavering support during stressful times and for sharing in the joy of celebrations. Her presence has been a source of great comfort and happiness.

IV. APPENDIX 93

I. Introduction

1. Introduction

1.1 DNA-dependent RNA polymerases in the tree of life

Gene expression is a fundamental cellular process, crucial for the proper functioning of myriad biological systems within the cell. The key to understanding this complex process lies in examining its initial step: Transcription. The primary role of DNA-dependent RNA polymerases (RNAPs) is to carry out this process, yet a detailed understanding of their structure is essential to fully comprehend their function [1]. Notably, the structural complexity of RNAPs varies across the three domains of life. While bacterial RNAP consists of five subunits, serving as a reference point, archaeal and eukaryotic RNAPs possess homologous subunits indicative of a shared ancestral origin [2]. These multimeric enzymes can be categorized based on the functions of their subunits: the catalytic core, the assembly platform, and auxiliary specialized functions [3]. In this section we will shortly analyze the structures of each polymerase (see figure 1.1), focusing on the factors that enable them to recognize and transcribe target genes.

1.1.1 Bacterial RNA polymerase

The discovery of Escherichia coli bacterial RNA polymerase (RNAP) in the 1960s marked a paradigm shift in our understanding of transcriptional processes, which laid the foundation for studying how a DNA molecule is intricately transcribed into its RNA counterpart [4]. In the late 80's, the first attempt to obtain a three-dimensional structure of this holoenzyme, RNAP bound to the σ_{70} regulatory subunit, was made by growing 2D crystals on positively charged lipid layers, negatively staining them, and then collecting electron microscopy data from various angles [5]. Although other RNAP structures, typically from the Thermus genus, have been available since then, it is only recently that a high-resolution structure of the E. coli RNAP bound to σ_{70} was obtained using X-ray crystallography [6].

This RNAP is an enzyme composed of multiple subunits: two copies of α, β, β', and ω, which then requires a σ factor to identify the promoter regions of the genes targeted for transcription [7]. Group 1 σ factors are composed of four flexibly connected regions: $\sigma_{1.1}$, σ_2 (conserved region 1.2-2.4), σ_3 (conserved region 3.0-3.1), and σ_4 (conserved region 4.1-4.2), with σ_2, σ_3, and σ_4 forming the structural domains [8]; of these, the σ_2 domain binds to the -10 element and the σ_4 domain to the -35 element [9].

1.1.2 Archeal RNA polymerase

In Archaea, like bacteria and unlike Eukarya, transcription is conducted by a single multimeric enzyme, yet their macromolecular machineries for replication, transcription, and translation more closely resemble those of Eukaryotic organism [3]. It has been shown that archaeal RNAP, with a total molecular weight of 370 kDa, possess homolog subunits to all but two RNAP II subunits: RPB8 and RPB9 [3]. This similarity was revealed through high-resolution structural analysis of the archaeal 13-subunit RNAP using X-ray crystallography. Notably, an exclusive subunit, Rpo13, was identified in Archaea,

distinct from pol II, suggesting a reason why fewer transcription factors are needed compared to Pol II [10]. It is established that archaeal RNAP requires only two transcription factors for transcription initiation: the archaeal TATA-binding protein (TBP) and transcription factor B (TFB), an ortholog of TFIIB [11].

In the archaeal transcription mechanism, TBP binds to the promoter region known as TATA box. This TBP-DNA complex is then recognized by TFB's C-terminus, initiating transcription 25 bp downstream of the TATA box [12]. TFB is known to specifically interact with a promoter region, the purine-rich motif called BRE (transcription factor B Recognition Element), located immediately upstream of the TATA box, which contributes along with TBP at determining the promoter strength [13]. Further, in the TFB-TBP-DNA complex, TFB was found to confer directionality to the transcription process [14].
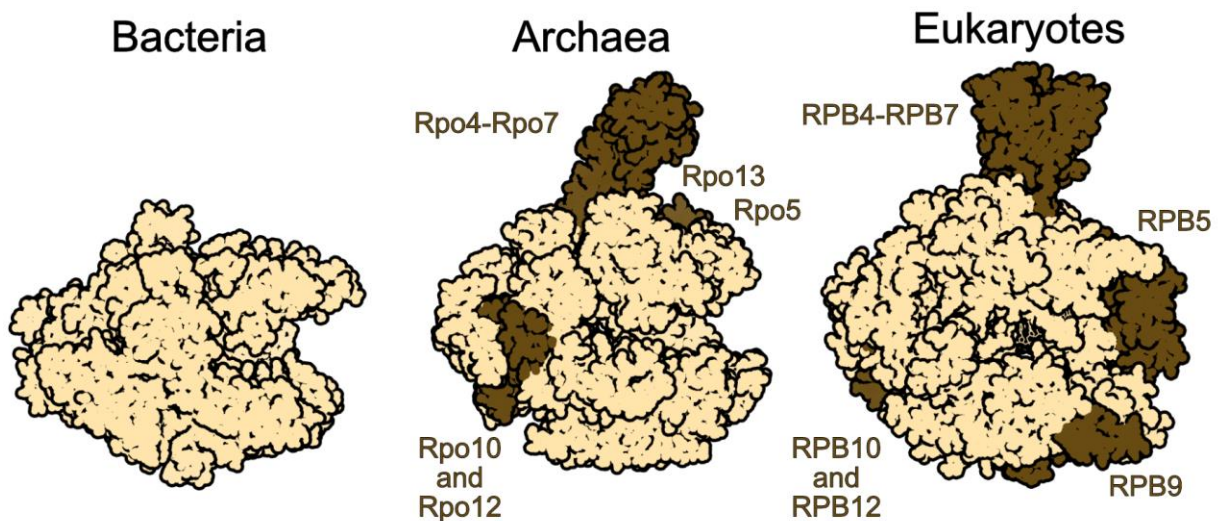
Figure 1.1 Comparison between RNAP in bacteria, Archaea and Eukaryotes (pol II), highlighting subunits unique to Archaea and Eukaryotes, which are absent in Bacteria. Beige-colored subunits represent universal conservation, while those in brown are exclusive to Archaea and Eukaryotes. PDBIDs: Bacteria: 1l6V, Archaea: 2WAQ, Eukaryotes: 1Y1W. Adapted from reference [2].

1.1.3 Eukaryotic RNA polymerase: Pol I and Pol II

There are three types of RNA polymerases, composed of multiple subunits, which transcribe all the nuclear genes in eukaryotic cells. RNA polymerase I (Pol I) synthesizes ribosomal RNA and RNA polymerase II (Pol II) produces messenger RNAs and small nuclear RNAs [15].

In yeast, Pol I is a multimeric enzyme with a molecular weight of approximately 589 kDa, comprising 14 subunits. It is responsible for transcribing the precursor rRNA gene, which is processed into 25S, 18S, and 5.8S rRNAs. The high-resolution structure of yeast Pol I, revealed through X-ray crystallography, displays distinct features, including a notably wider DNA-binding cleft [16]. More recently, the structure of human Pol I was resolved using cryo-EM. It was found to consist of 13 subunits, unlike the 14 in yeast, attributed to its stalk being composed of a single subunit as opposed to two in yeast [17]. This study also successfully captured the human Pol I-RRN3 complex, highlighting the interaction of this initiation factor

with the Pol I stalk. Refer to Figure 1.2 A for a detailed visualization and further information regarding the general transcription factors associated with Pol I.

Pol II is the pivotal multimeric enzyme responsible for transcribing genes encoding proteins in all eukaryotic cells [18]. Remarkably, the structure of RNA Polymerase II was the first among polymerases to be resolved at high resolution, a feat achieved at the beginning of this century [19]. Comprising 10 subunits, this enzyme reveals critical structural elements organized into four mobile modules: core, jaw-lobe, shelf, and clamp, each integral to its function. Furthering our understanding, the structure of the 12-subunit bovine Pol II was elucidated using cryo-EM [18]. The structural differences between these variants shed light on their conformational control during enzymatic activity. Notably, Pol II interacts with an array of complex transcription factors, with one such factor, TFIID, reaching a size of 1.3 MegaDaltons [20]. For a more detailed view of the transcription factors assisting Pol II in initiating transcription, please see Figure 1.2 B.
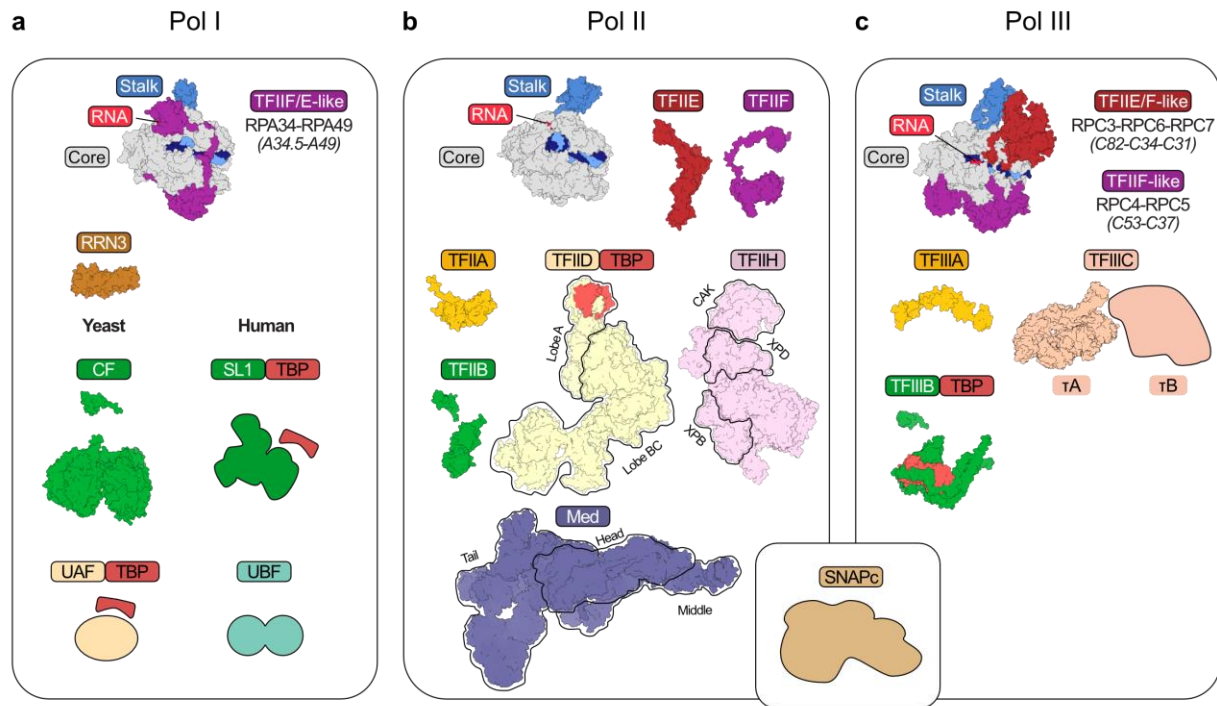
Figure 1.2. General architecture of RNA polymerases I, II, and III, along with their specific general transcription factors. A. Structure of Human Polymerase I elongation complex, highlighting transcription factors critical for transcription in both yeast and human. B. Structure of Bovine Polymerase II elongation complex and its essential transcription factors in human. C. Structure of Human Polymerase III elongation complex showing its yeast transcription factors. Transcription factors in surface representation indicate that their structures have been resolved. This figure was adapted from reference [20].

1.2 RNA polymerase III transcription

The transcription of small non-coding RNAs, such as, 5S rRNA, tRNA, VAI and U6 small nuclear RNA is carried out by RNA polymerase III [21], [22]. The transcription of these genes requires not only the presence of Pol III, but also promoter elements and auxiliary transcription factors [23]. The genes

transcribed by Pol III are called class III genes and they are characterized by the presence of three types of promoters within the gene (type I and II) or outside the gene (type III) (see Figure 1.XXXX) [24]. The regions in the DNA sequence recognized by specific transcription factors and located inside the gene are called internal control regions (ICRs). ICRs are highly conserved and discontinuous DNA segments, which are classified, depending on their consensus sequences, in: A Box , B Box and C Box [25]. These conserved sequences serve as binding platforms for specific transcription factors. In the case of tRNA, the transcription factor (TF) IIIC binds to A Box and B Box, while the presence of the C Box in 5S rRNA genes allows the binding of transcription factor (TF) IIIA [26].

In the upcoming section, I will explore in detail the various types of Pol III promoters, exemplified by 5S rRNA for Type I, tRNA for Type II, and U6 snRNA for Type II.

1.2.1 Transcription from type I promoter
The isolation and identification of a 37 kDa polypeptide from Xenopus laevis ovarian extracts were achieved due to its unique ability to bind to the Internal Control Region (ICR) of a 5S gene in vitro, facilitating its transcription [27]. The Internal Control Region (ICR) of 5S genes comprises three discrete elements: an A-box with low TFIIIA affinity, an intermediate element, and a C-box, both exhibiting higher affinity towards TFIIIA (see figure 1.3 - top) [28]. TFIIIA, a protein of approximately 40 kDa, was found to contain 13 cysteine residues that coordinate zinc ions, a feature critical for its binding to the promoter sequence [29]. These cysteines are part of 9 zinc fingers, with the first three N-terminal fingers essential for DNA binding, and fingers 4-7 significant for RNA binding [30]. It has been shown that the order of transcription factor recruitment is as follows: Initially, promoter recognition is achieved by TFIIIA, which then recruits TFIIIC, a multisubunit protein also vital for type 2 promoters (refer to section 1.3 for its composition and function). Following the interaction of these two complexes, TFIIIB is recruited, subsequently leading to the recruitment of Pol III. [31].

1.2.2 Transcription from type II promoter
Similar to the Type I 5S promoter's non-continuous intragenic structure, tRNA genes exhibit two distinct regions, named A- and B-blocks. These blocks harbor conserved sequences that correspond to the D-loop and T-stem and T-loop of the tRNA structure, respectively [32]. The segment between these promoter regions has been shown to be non-essential for transcription, as point mutations in this area do not disrupt transcriptional activity [28]. However, the length of this intervening region influences the affinity of TFIIIC for these intragenic promoters [33], with variations in nucleotide distance ranging from 31 to 93 bp [23]. Initially, a transcription factor termed factor τ, now identified as TFIIIC, was observed to bind these motifs in tRNA without assistance from other transcription factors [34]. TFIIIC's interaction is predominantly directed by the B-box, demonstrating nanomolar (nM) affinity [35], while the A-box exhibits a comparatively weaker micromolar (μM) affinity [36], [37]. Upon tRNA promoter recognition, TFIIIC facilitates the recruitment of TFIIIB, which then polymerase III binds to it (see figure 1.3 -middle) [38].

1.2.3 Transcription from type III promoter
The principal distinction between Type I and II promoters, compared to Type III promoters, lies in the positioning of the motif region. In Type III promoters, the motif region is situated externally to the gene

4

targeted for transcription [39]. Notably, the U6 RNA gene from Xenopus tropicalis, which demonstrates resistance to full inhibition by α-amanitin and possesses a TATA box and proximal elements outside the gene, was one of the initial Type III genes identified to be transcribed by polymerase III [40]. Initially, it was revealed that transcription of this gene type necessitated at least four transcription factors, including TFIIIB, while TFIIIA and TFIIIC were deemed unnecessary [41]. Subsequent research clarified that, in addition to TFIIIB binding to the TATA box, another multisubunit complex was essential. This complex was variably named by different research groups, either as the PSE transcription factor (PTF) [42] or the snRNA activating protein complex (SNAPc) [43], but it was eventually established that both names referred to the same entity (see figure 1.3 - bottom).
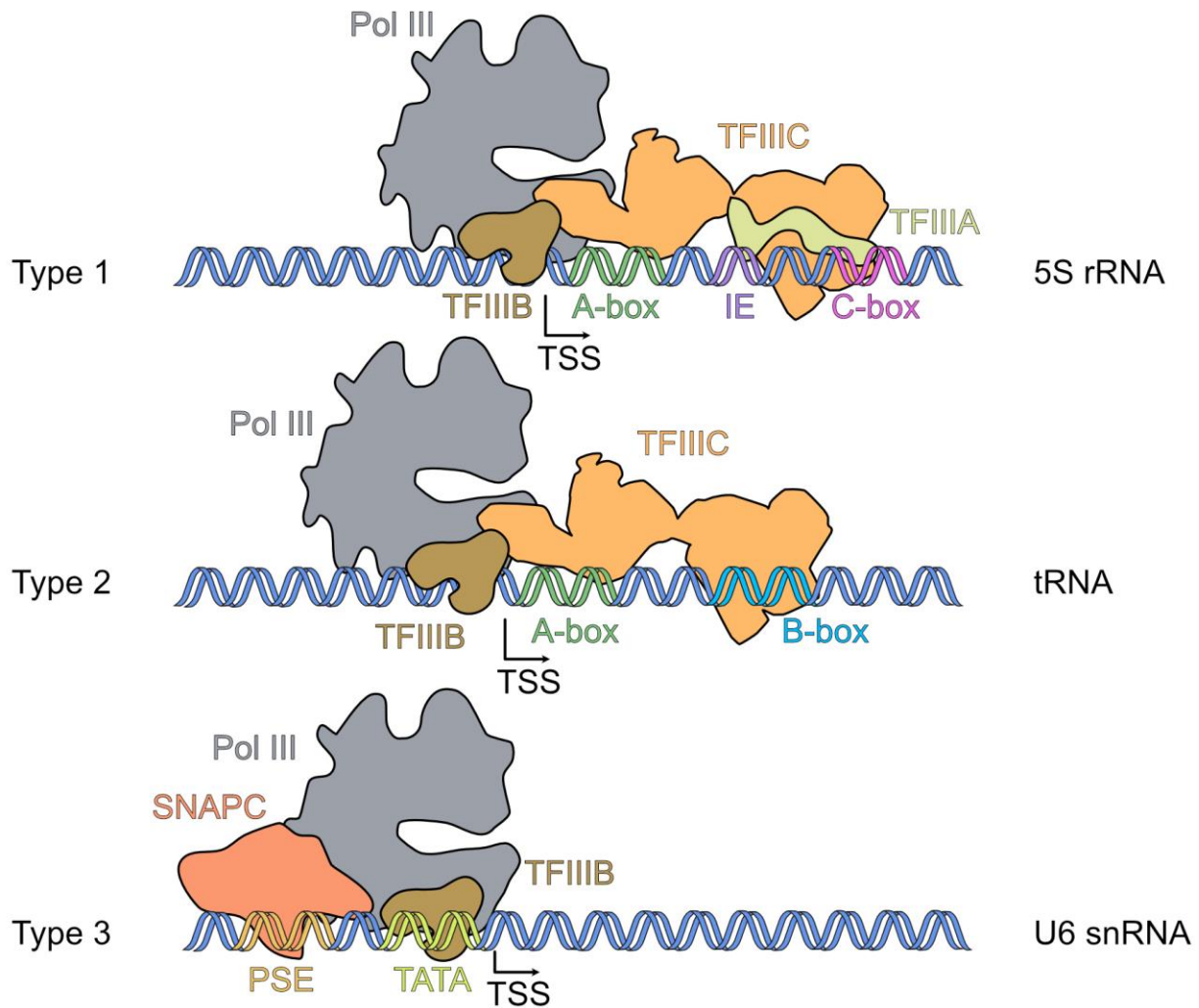
Figure 1.3 Different types of pol III genes. At the top, the type 1 gene, represented by 5S rRNA, displays its three promoter regions: A-box, intermediate element (IE), and C-box, along with the transcription factors necessary for transcription, namely TFIIIA, TFIIIC, and TFIIIB. In the middle, tRNA gene, a type 2 promoter, characterized by the intragenic promoters A-box and B-box, requires only the transcription factors TFIIIC and TFIIIB for transcription. At the bottom, the type 3 gene, exemplified by U6 snRNA, features external promoter regions - the Proximal sequence element (PSE) and TATA box - and its transcription factors SNAPC and TFIIIB for transcription. TSS: Transcription Start Site. Adapted from reference [20].

5

1.3 Transcription factor IIIC (TFIIIC)

1.3.1 Discovery of TFIIIC in yeast

Traditionally, the study of transcription mechanisms in eukaryotic cells relied heavily on the development of crude in vitro transcription systems [34]. These cell-free extracts played a pivotal role in understanding the functions of various transcription factors and even led to the discovery of previously unknown factors involved in transcription. However, this methodology had a significant limitation: it did not permit the detailed analysis of individual components' roles or elucidate the significance of specific transcription factors within the broader transcription system under study. To address these limitations, researchers began with preliminary fractionation of cell extracts to identify potential transcription factors. For instance, to understand the role of polymerase III in tRNA transcription, it was essential to isolate the transcription factors involved in promoter recognition.

The first attempt to isolate the Transcription Factor IIIC (TFIIIC) in Saccharomyces cerevisiae led to the partial purification of a factor initially known as the factor τ. This factor was distinguished by its stable interaction with tRNA, independently of RNA polymerase III and other transcription factors. The factor τ was reported to have a molecular weight of approximately 300,000 kDa, determined through glycerol gradient sedimentation, and it was suggested its multimeric nature [34]. Subsequent research revealed that the τ factor protected around 90 base pairs of the $tRNA_3^{Glu}$ gene (-11 to +88) as shown in DNase I footprinting experiments. These studies also indicated that the factor τ had a weaker binding affinity for the A-box compared to the B-box, as confirmed by λ exonuclease assays. This was further corroborated by dimethyl sulphate (DMS) protection experiments, which demonstrated more significant contact points of the factor with the DNA at the B-box, fewer at the A-box, and none in the intervening regions [44]. In another study, the equilibrium constant of TFIIIC was determined to be 1 x 10^{10} M^{-1} (or Kd of 0.1 nM) using Electrophoretic Mobility Shift Assay (EMSA). This study found that mutations in the B-box significantly impacted the equilibrium constant, affecting it by up to 370-fold and reducing TFIIIC's interaction with the tRNA more severely than mutations in the A-box, which at most resulted in a 5-fold reduction in binding [45].

Following the isolation of factor τ, several critical questions remained unanswered, shaping the direction of subsequent research. These questions included: (1) determining the exact number of proteins that constitute factor τ; (2) ascertaining whether the binding to the A- and B-boxes is independent; (3) investigating if factor τ binds exclusively to the A- and B-boxes, or if other regions of the tRNA gene are also crucial for binding; and (4) understanding how transcription proceeds if factor τ is stably bound to the gene [44].

To address the outlined questions, it was essential to first understand the composition of the factor τ. Proteolytic cleavage was employed as a method to dissect its structure and find the DNA-binding domains. This approach led to the identification of a protease-resistant domain, designated as τB, which specifically binds to the B-Box. Another domain, termed τA, was found to interact with the A-Box. A flexible hinge was suggested to link these two domains, highlighting a key structural relationship within the τ factor. [46]. Further insights into the molecular structure and arrangement of these τA and τB domains were gained through scanning transmission electron microscopy studies. These studies

revealed two closely associated globular domains, each approximately 300 kDa, in the absence of DNA. Interestingly, when DNA with varying distances between the A- and B-boxes was introduced, the complex adopted different conformations. A dumbbell-shaped structure was observed, corresponding to the τA and τB domains, when the distance between these boxes was large. In contrast, when this distance was short, the complex's structure became unresolved. This suggests the existence of a hinge region that connects the τA and τB domains, allowing for conformational flexibility in response to the spacing between the A- and B-boxes [47]. An additional study aimed at understanding the effects of varying distances between the A-Box and B-Box motifs on TFIIIC binding. In this study, the distance was systematically reduced from 74 base pairs (bp) in the wild type (wt) to as little as 0 bp, including oligonucleotides of intermediate lengths. The results indicated that the distance between these motifs significantly influences the interaction, particularly affecting τA binding. An optimal distance of 53 bp for effective binding was observed [33].

To elucidate the mechanism of tRNA gene activation, a key objective was to identify the components of TFIIIC and the polypeptides that interact with DNA. This investigation began with a study that improved the general protocol for purifying TFIIIC. Following this, TFIIIC was incubated with tRNA genes, and the resultant TFIIIC-DNA complex was isolated using Electrophoretic Mobility Shift Assay (EMSA) followed by SDS-PAGE for subunit identification. These experiments revealed four major components with apparent molecular weights of 145 kDa, 135 kDa, 100 kDa, and 65 kDa, which aligned with findings from a previous study [44]. Antibodies generated against these components showed that those binding to the 145 kDa and 100 kDa polypeptides affected the migration of the TFIIIC-DNA complex in EMSA and partially inhibited transcription. The same antibodies led to the identification of the 145 kDa component as part of the τB domain. UV cross-linking experiments further confirmed the interaction of the 145 and 100 kDa bands with DNA [48]. In an additional study employing a novel, longer photocrosslinker, not only were the 145 and 100 kDa subunits confirmed to bind to DNA, but the previously mentioned 135 kDa and 55 kDa subunits were also identified bound to the DNA [49]. These discoveries were particularly significant as it contributed to identifying potential components involved in DNA-protein interactions within TFIIIC.

Subsequent research efforts successfully identified and heterologously cloned the polypeptides previously described within TFIIIC. Notably, the 100 kDa polypeptide, now known as τ95, was the first to be cloned and peptide sequenced. This accomplishment was a key factor in naming its gene TFC1 (Transcription Factor C 1). The τ95 subunit consists of 649 amino acids and has a theoretical molecular weight of 73,509. It was predicted to contain helix-turn-helix structures and a highly acidic region. However, studies revealed that τ95, despite having this helix-turn-helix domain, did not exhibit DNA-binding properties when expressed heterologously. This finding highlighted the importance of other TFIIIC subunits in facilitating DNA binding.[50], [51]. Following this, the TFC3 gene encoding τ138, a B-box binding subunit, was cloned. τ138 is a 1160 amino acid protein with a molecular weight of 132,026 kDa. Predictions indicated the presence of High Mobility Group (HMG) domains at both its N-terminus and C-terminus. Like τ95, τ138 did not exhibit DNA-binding activity in its recombinant form [52].

After isolating the two subunits found to bind the A- and B-box, the TFC4 gene was cloned. This gene corresponds to the τ131 subunit, a protein with a theoretical molecular weight of 120,153 and

7

consisting of 1025 amino acids. Sequence analysis of τ131 revealed the presence of 11 tetratricopeptide repeats distributed throughout the protein: TPR 1-5 in the N-terminus, TPR 6-9 in the middle, and TPR 10-11 in the C-terminus. Moreover, a Basic helix-loop-helix motif was identified between TPR 6-9 and TPR 10-11. τ131 was hypothesized to serve as a hinge region, providing TFIIIC with its flexibility and potentially facilitating interaction with TFIIIB. [53]. Subsequently, the gene encoding the 91 kDa protein was cloned and named TFC6. This protein with a predicted molecular weight of 74 kDa is characterized by a highly acidic N-terminal region and a cluster of 13 cysteines capable of potentially binding zinc. Notably, τ91 was observed to enhance the affinity of mutants in the τ138 subunit of the TFIIIC-DNA complex. This was evidenced by the identification of a suppressor mutation (E330K) in τ91. However, in the same study, when specific antibodies against τ91 were used, they did not succeed in localizing τ91 within the protease-resistant τB-DNA complex [23].

The two smallest, and last, subunits of TFIIIC, historically less understood, include those encoded by the TFC7 gene. TFC7 produces a chimeric protein comprising 435 amino acids with a theoretical molecular weight of 49 kDa. The N-terminus of this protein is related to acid phosphatase, while its C-terminus is vital for interaction with τ95. The discovery of a stable τ55-τ95 complex, distinct from the larger TFIIIC complex, led to the hypothesis of an alternate function for τ55. This function is likely associated with bridging transcription and metabolism, as indicated by findings that a mutation in the N-terminus of τ55 did not affect the affinity of TFIIIC for DNA but did result in a growth defect upon a change in carbon source [54]. τ60, the last subunit of TFIIIC to be characterized, consists of 588 residues and has a molecular weight of 67,640. It is located, at least partially, within the τB domain and is believed to have a role in recruiting TBP (TATA-binding protein). This is supported by its direct interaction with TBP, estimated to have an affinity of 100 nM [55].

The distinction between subunits comprising the τA or τB domains of TFIIIC has not always been clear. The composition of τA has been somewhat clear regarding certain subunits, but remained ambiguous for others. For instance, electron microscopy studies utilizing a β-galactosidase-τ95 fusion protein helped locate τ95 within the τA domain, which is responsible for A-box interaction [56]. This also confirmed the presence of one subunit in the TFIIIC complex, resolving a debate over the existence of two τ95 subunits in a single TFIIIC complex [49]. UV cross-linking experiments showed that τ95 subunit was located close to the A-box, along with the τ55 subunit, and τ131 was weakly cross-linked to the A-box [49]. An additional study suggested that a 100 kDa protein (now identified as τ95) was able to bind to the DNA [48].

It is important to note that τB as observed through electron microscopy and τB derived from limited proteolysis should not be assumed to be equivalent [55]. Regarding the τB subcomplex, UV cross-linking experiments conclusively showed that τ138 is associated with the B-box [48], [49]. While τ91 was shown to aid τ138 in DNA binding, initial attempts failed to localize it in the B-box binding region [23]. Conversely, the τ60 subunit, at least its C-terminal region, was located in τB. Additionally, τ60 was suggested to act as a bridge between the τA and τB subcomplexes [55], [57].

Following the successful identification and cloning of all subunits, efforts were made to reconstitute independently the minimal recombinant τA (rτA) and τB (rτB) domains in insect cells. The minimal rτB

was assembled using τ138, τ91, and τ60 subunits, forming a complex with strong DNA-binding capabilities. This module had similar binding properties as the previously described protease-resistant τB. Among these, τ91 was the only subunit that demonstrated DNA-binding properties when isolated. In contrast, the minimal rτA was reconstituted with τ131, τ95, and τ55 subunits. However, this subcomplex exhibited non-specific DNA-binding properties, which were partly attributed to the τ95 subunit. This non-specific binding observed in rτA raised questions about its binding specificity [58].

1.3.2 Structural studies of yeast TFIIIC

With the molecular composition of yeast TFIIIC now elucidated and the minimal rτA and rτB modules successfully reconstituted [58], there still exists a gap in structural studies. These studies are essential to further elucidate the complex properties of this intricate transcription factor, particularly its potential role in the initiation of RNA polymerase III transcription.

X-ray crystallography provided groundbreaking insights into the core of the τB subcomplex, particularly regarding the structures of the τ60 and Δτ91 (a truncated version of τ91 lacking the N-terminal part) subunits. Prior to this, only sequence-based predictions were available. τ60 was found to have a seven-blade β-propeller structure at its N-terminus and a novel α/β fold at its C-terminus. Similarly, the Δτ91 subunit featured an N-terminal extension attached to a C-terminal β-propeller. Intriguingly, these β-propellers from τ60 and Δτ91 interacted perpendicularly, with Δτ91's propeller packed against that of τ60. This interaction was suggested to form the core of the τB subcomplex, which notably did not bind DNA. Instead, it was proposed that this core serves as a scaffold for τ138, the subunit capable of DNA binding [59]. Regarding the τ138 subunit, an extended winged helix (eWH) domain, spanning amino acids 546 to 641, was crystallized. This eWH domain exhibited weak and non-specific binding to both single-stranded and double-stranded DNA [60].

In comparison with the τB subcomplex, the τA subcomplex of TFIIIC has been the subject of more extensive structural studies. The first crystal structure solved within this subcomplex was the phosphatase domain of the τ55 subunit in *Schizosaccharomyces cerevisiae*. This structure exhibited a canonical histidine phosphatase fold, consisting of a core formed by 6 β-strands (arranged in both parallel and antiparallel configurations) surrounded by 9 α-helices. The enzymatic activity of this domain was found to primarily influence metabolic processes [61].

In studies involving *Schizosaccharomyces pombe*, the dimerization domain of the sfc1/sfc7 heterodimer (corresponding to the τ55/τ95 heterodimer in *Schizosaccharomyces cerevisiae*) was crystallized. The resulting structure revealed a complex interaction among 3 β-barrels. Additionally, the C-terminal part of sfc1 (equivalent to τ95 in *Schizosaccharomyces cerevisiae*), known as the DNA binding domain (DBD), was also crystallized and its structure determined. This structure comprised a winged helix domain followed by a novel arrangement of 5 α-helices and 2 β-strands, displaying micromolar affinity for both single-stranded and double-stranded DNA [62].

The τ131 subunit was also studied; its N-terminal TPR domain (amino acids 123 – 566) was solved using X-ray crystallography. It featured a total of 10 TPR domain repeats divided into two 'arms' by an extended helix and a 'ring' domain. This structure demonstrated binding to BDP1 and a small region of

the τ138 subunit known as τIR (τ131 interaction region), with a high affinity in the nanomolar range (80 - 100 nM). A longer version of the τ131 crystallized protein (amino acids 1 − 566) showed interactions with the other two subunits of TFIIIB: brf1-TBP [60].

While crystallography provided valuable insights into the functioning of TFIIIC, many questions remained unanswered by this 'divide and conquer' approach. A different strategy to determine the molecular architecture of TFIIIC involved cross-linking mass spectrometry. This technique, combined with available crystal structures led to the most updated structural information on TFIIIC. Additionally, mutagenesis and biophysical methods were key in identifying the binding regions of τ131 and τ138, as well as the interactions between τ131 and TFIIIB components [60].

The 'resolution revolution' in cryo-electron microscopy has enabled high-resolution analysis of large, flexible complexes [63], a development that has significantly advanced the study of transcription factors [64]. Utilizing this technique, the structure of the minimal τA subcomplex of TFIIIC was resolved. This revealed similarities with structures previously solved via X-ray crystallography and provided new insights into the overall architecture of the τA subcomplex. For example, the positioning of the DBD-τ95 within the structure suggested an unconventional mode of DNA binding. Furthermore, combining negative staining and cross-linking spectroscopy helped to ascertain the overall positioning of the Brf1-TBP complex in interaction with the τA subcomplex [37].

More recently, the structure of the TFIIIA-TFIIIC-Brf1-TBP complex bound to a 5S rRNA gene was elucidated. This study not only detailed the complex interactions among these different transcription factors and their engagement with DNA but also shed light on the overall structure of TFIIIC. For instance, it delineated the distribution of TFIIIC's largest subunit, τ138, across the τA and τB subcomplexes, with the C-terminus allocated to the former and the N-terminus to the latter. The presence of the Brf1-TBP complex was observed to alter (or at least stabilize) the interaction of the τ131 subunit with DNA [65]. These structural findings of TFIIIC align in some aspects but differ in others when compared to TFIIIC bound to a type 2 pol III gene [66] . This comparison and its implications will be discussed in section 3.3.

1.3.3 Understanding the subunit composition of TFIIIC in human
Unlike yeast TFIIIC, whose activity was linked to a single fraction (refer to section 1.3.2 for details on the discovery of TFIIIC), identifying the polypeptide composition of human TFIIIC proved to be more challenging and, at times, yielded contradictory results.

Initial efforts to decipher the composition of human TFIIIC, utilizing FPLC Mono Q chromatography, identified two functional components: TFIIIC1 and TFIIIC2. TFIIIC2 was responsible for B-box recognition, while TFIIIC1 complemented it for in vitro transcription of tRNA and VA RNA genes. DNase I footprinting experiments demonstrated that increasing concentrations of TFIIIC1 expanded the protective role of TFIIIC2 to include the A-Box, though TFIIIC1 alone did not protect the B-box sequence. Sedimentation analysis further revealed that TFIIIC2 had a molecular weight of approximately 400-500 kDa, whereas TFIIIC1 was around 200 kDa [67]. An independent study using DNA affinity chromatography, instead of

Mono Q chromatography, confirmed the separation of TFIIIC in TFIIIC1 and TFIIIC2 components and that both are need for the transcription of tRNA and VA RNA genes [68].

Subsequent research, however, presented conflicting results. One study reported that human TFIIIC in solution consisted of a single globular protein of approximately 120 kDa [69], contrasting with earlier findings. For instance, a study determining the composition of TFIIIC2 via SDS-PAGE identified at least 10 bands ranging from 45 kDa to 250 kDa [67] and another showed polypeptides in purified TFIIIC2 samples of 230, 110, 100, 80, and 60 kDa [70]. This multimeric composition was later supported by another study, which found TFIIIC2 composed of similar-sized polypeptides (220, 110, 102, 90, and 63 kDa), with the 220 kDa subunit crosslinked to the VA1 promoter. Interestingly, this study also noted two forms of TFIIIC2 (upper band and lower band form), differing mainly in the presence or modification of the 110 kDa subunit in the lower form and its transcription activity [71].

After numerous studies confirmed the multimeric nature of the TFIIIC2 component, research efforts shifted towards identifying, cloning, and characterizing the genes encoding its subunits. The first major breakthrough was the identification of the gene for the largest subunit of human TFIIIC, a 230 kDa protein, which shared limited similarity with the 138 kDa yeast counterpart [72]. Subsequently, the gene for the 110 kDa subunit was identified and cloned. This subunit, which differed in the composition of two types of TFIIIC2 [71], encodes a 911 amino acid protein with a molecular weight of 100.7 KDa. The protein was predicted to have acidic and basic regions in its N-terminus and WD40 repeats in the C-terminal part [73].

Further progress led to the identification of two proteins in a single study: hTFIIIC63 and hTFIIIC102. The gene encoding hTFIIIC63 produced a 60 kDa protein (519 amino acids) with a 22% sequence identity to τ95. On the other hand, hTFIIIC102, an 886 amino acid protein weighing 101 kDa, shared a 31% identity with τ131 and displayed TPR repeats throughout its sequence [74].

The last polypeptide among the original five bands thought to be the sole components of TFIIIC2 was hTFIIIC90. The gene for this 822 amino acid (92 kDa) protein showed no similarity to any yeast TFIIIC subunit. However, it was crucial for RNA polymerase III transcription and exhibited unexpected acetyl-transferase activity. Moreover, hTFIIIC90 was identified as interacting with hTFIIIC220, hTFIIIC110, and hTFIIIC63 [75].

It took nearly a decade to identify the final missing subunit of human TFIIIC, which had been overlooked during endogenous purification. This subunit, named TFIIIC35 and consisting of only 215 amino acids, was related to yeast's smallest subunit, τ55. It was found to interact with hTFIIIC63 (τ95), resolving a longstanding debate about the exact composition of human TFIIIC. This discovery revealed that human TFIIIC, like yeast TFIIIC, comprises the same number of subunits [76]. A comparison of the molecular weights and homology between yTFIIIC and hTFIIIC is illustrated in Figure 1.4. The low percentage of identity between the subunits of these two species elucidates the challenges in identifying orthologs across different organisms [77]. In parallel, the composition of TFIIIC1 also posed difficulties. However, at least one component was clearly identified: the human Bdp1 [78].
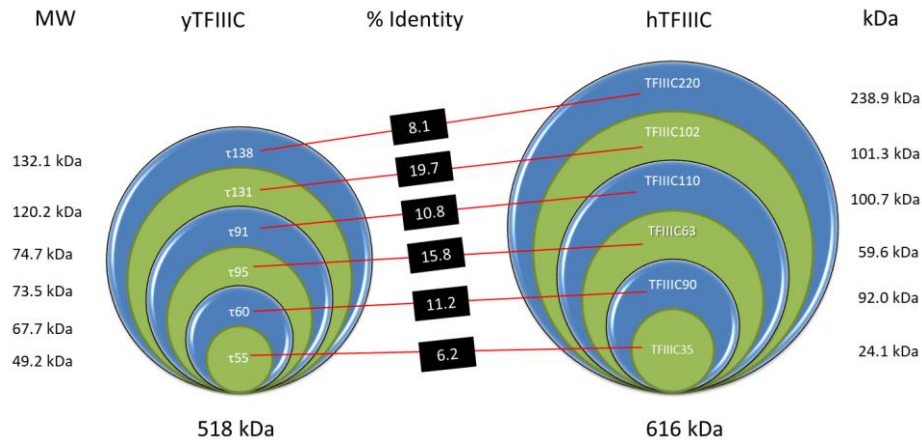
Figure 1.4: Limited sequence conservation between yeast and human TFIIIC subunits. yTFIIIC stands for yeast TFIIIC and hTFIIIC for human TFIIIC. Source: Adapted from [77].

1.4 Transcription factor IIIB

The TFIIIB from yeast to human is well conserved. This complex harbors there subunits: TATA-binding protein (TBP), B double prime (Bdp1), and B-related factor 1 (Brf1) [79]. In vertebrates, there are two isoforms of the last subunit: Brf1 (TFIIIB-β) or Brf2 (TFIIIB-α). Brf2 has been found to participate in the regulation of cellular processes, such as oxidative stress [80] .

TFIIIC is considered as an assembly factor, because it positions TFIIIB upstream of the transcription start site. TFIIIC is dispensable once the transcription starts, and it is TFIIIB, alone, which positions Pol III for repeated transcription cycles [81].

Different groups have been focusing on biophysical and biochemical studies of TFIIIB to understand its mechanism when it binds to the DNA. Just recently, the crystal structure of the human TFIIIB-DNA complex was solved to gain insight into the structure and function of Bdp1 [82]. Additionally, the structure of Pol III-TFIIIB complex in *Saccharomyces cerevisiae* solved by two different groups helped to understand how TFIIIB recruits Pol III [83], [84].

1.5 TFIIIC as genome organizer

Two initial genome-wide studies in *Saccharomyces cerevisiae* provided insights into the occupancy of the Pol III transcription machinery under various environmental and cellular conditions. The first study revealed that under conditions of nutrient deprivation, TFIIIC occupancy within the genome significantly increased. The second study found that TFIIIC remained bound to its target sites during the late growth phase. A common finding in both studies was a notable decrease in Pol III occupancy under these conditions [85], [86]. However, these studies did not identify specific genomic regions where TFIIIC was present independently, without the accompanying TFIIIB and Pol III.

A subsequent study using a similar genome-wide approach revealed an unexpected finding: eight loci, termed ETC (Extra TFIIIC sites), showed exclusive occupancy by TFIIIC without the presence of Bdp1, Brf1 (TFIIIB components), or Rpc34 (a Pol III component). These ETC loci contained a conserved sequence with a B-box and 10 additional nucleotides, three of which were highly conserved. Additionally it was

shown that these sites were also conserved in a closely related yeast species. Despite the identification of these loci, their function remained unclear [87].

A subsequent study in fission yeast began with a high-resolution ChIP analysis of Inverted Repeats (IR) elements, recognized as boundary elements in the genome. This analysis revealed that IR elements contain multiple B-box sequences that recruit TFIIIC, but notably, Pol III was absent at these sites. The presence of TFIIIC at these IR elements was crucial for preventing the spread of heterochromatin, underscoring its role in maintaining genomic integrity. A further examination beyond IR elements discovered that TFIIIC was also highly concentrated at other genomic loci termed Chromosome-Organizing Clamps (COC). These COC loci, like IR elements, lacked Pol III presence and were found to be associated with the nuclear periphery. This association suggests a role for TFIIIC in facilitating higher-order chromosomal organization, potentially through loop formation in the fission yeast genome [88]. Building on these findings, a later study in budding yeast confirmed that six of the eight previously identified Extra TFIIIC (ETC) sites were also located at the nuclear periphery, further highlighting the conserved role of TFIIIC in genome organization across yeast species [89].

In an effort to understand the role of TFIIIC as a genome organizer beyond yeast and in metazoans, a genome-wide study on human cells identified approximately 1,865 Extra TFIIIC (ETC) sites. These sites were characterized by either an extended B-box or a novel ETC motif. A significant correlation was observed between the presence of TFIIIC and CTCF, a protein involved in chromosome organization with cohesin, at these ETC sites [90]. This correlation between TFIIIC and CTCF, as well as between cohesin and TFIIIC, was also confirmed in the mouse genome [91]. However, another study attempting to verify this interaction through coimmunoprecipitation (co-IP) experiments did not yield conclusive results. Instead, a new interaction with the condensin II complex was identified and confirmed using proximity-dependent biotin identification (BioID) techniques. ChIP sequencing further validated these findings, showing strong colocalization of TFIIIC220 with two subunits of condensin II: NCAPH2 and NCAPD3 [92]. More recent research has investigated the role of human TFIIIC in chromatin looping. This study revealed that during stress conditions, hTFIIIC interacts with Alu elements (AE) and controls gene expression through its ability to form chromatin loops and through histone acetylation activity[93]. Notably, the capacity of hTFIIIC to form loops was shown to vary depending on the cell type being analyzed [94]

These discoveries over the past two decades have highlighted TFIIIC's remarkable versatility as a genome organizer, a role extending beyond its well-established function in transcription. This dual function of TFIIIC underscores the need for further structural studies in the coming years to better understand this fascinating complex.

1.6 Using cryo-electron microscopy to study macromolecular complexes

Electron microscopy has been a crucial instrument for several decades, with the first structure determined using this technique in the early 1990s [95]. It was not until 2017 that Richard Henderson, Joachim Frank, and Jacques Dubochet were awarded the Nobel Prize for their contributions to the development of cryo-Electron Microscopy [96]. While their contributions were fundamental to the advancement of this technique, it was the introduction of direct electron detectors [97] and the development of sophisticated software for image processing [98] that truly revolutionized the field.

1.6.1 Standard workflow for a cryo-EM study

After purifying your sample to homogeneity, the next critical step is preparing it for cryo-Electron Microscopy (cryo-EM) using a vitrobot. This step is essential as it safeguards the sample against the vacuum conditions within the microscope [99]. While this process might seem straightforward, it involves several key considerations, such as optimizing sample carriers and treatments, deposition techniques, and vitrification [100]. Working with a vitrobot presents its own set of challenges, particularly in achieving reproducible ice thickness – a crucial factor for successful imaging. Traditionally, the vitrobot utilizes a blotting paper method, whose mechanism is not fully understood. The issue of variable ice thickness, often non-reproducible, has been recently investigated, highlighting the influence of irregular blotting paper as a contributing factor [101]. Another challenge in sample preparation for cryo-Electron Microscopy (cryo-EM) is the air-water interface issue, where many proteins and complexes tend to interact strongly, leading to partial dissociation, denaturation, or even preferred orientation [102]. To mitigate these interactions, strategies such as the use of detergents [103] or different types of grids, like carbon-coated or graphene grids, can be employed [104]. These approaches help prevent, in most cases, the adverse interactions of the complex with the air-water interface.

Once your complex is stabilized for sample preparation, the next step is screening your prepared grid to identify the most suitable one – characterized by minimal aggregation and a higher visibility of particles. Following this, the data collection phase becomes crucial. This involves capturing thousands of movies of your cryo-immobilized sample. Typically, this data collection is automated and continuously monitored and adjusted to optimize the results [105].

After data collection, the next phase is image processing. This is where software like Relion or cryoSPARC comes into play, processing the images to transform 2D projections of the protein of interest into 3D volumes with high-resolution features. It involves several steps, including particle picking, 2D classification to identify and discard poor-quality particles, followed by 3D classification to sort particles into different conformational states. This is then followed by 3D refinement to improve the resolution of the structure, and finally, post-processing steps are applied to correct for the effects of the contrast transfer function (CTF) and to sharpen the map, which reveals detailed structural information.

1.7 Aims and scope of this thesis

I have shown from the discovery and subsequent investigation of yeast and human TFIIIC, tracing the progress from the initial identification and cloning of its subunits to advanced structural analyses. Initially, X-ray crystallography was employed to dissect the functions of distinct subunits and domains within TFIIIC. However, this approach was limited in its ability to elucidate the complete molecular structure of TFIIIC. The integration of crosslinking mass spectrometry and cryo-electron microscopy (cryo-EM) techniques facilitated a more comprehensive understanding of the overall molecular architecture of the full TFIIIC complex and the τA subcomplex in addition of the structure of the TFIIIA-TFIIIC-Brf1-TBP complex bound to other type of pol III gene.

Prior to this research, high-resolution structural data for the τB subcomplex and the entire TFIIIC complex were lacking, particularly regarding their interactions with DNA promoter regions of varying affinities. Also, information about the flexible linker between τA and τB was still missing. Utilizing cryo-EM, this thesis presents the first high-resolution structures of both τA and τB subcomplexes, achieving a significant breakthrough in understanding DNA binding affinity and specificity. These structural results were further supported by biochemical and biophysical characterizations.

In a similar approach, high-resolution cryo-EM structures of yeast TFIIIC bound to DNA were obtained, shedding light on the τA-DNA interaction and the roles of specific domains. To probe the dynamics of TFIIIC-DNA interactions, a tagged DBD of τA subcomplex was developed for use in single-molecule fluorescence microscopy experiments. These experiments revealed a dynamic interaction between TFIIIC and DNA, challenging previous perceptions and contributing new insights into the transcription factor's promotor recognition.

II. Results and discussion

2. Cryo-EM analysis of human TFIIIC

Given that prior research in our laboratory exclusively involved the yeast variant of Transcription Factor IIIC (TFIIIC), with no projects focusing on its human counterpart, my project represented a new direction for our research. This shift entailed a transition from yeast-based studies to exploring the structural aspects of the human TFIIIC complex.

The primary objective of my research project was to elucidate the complete structure of the TFIIIC complex and to revel its underlying structural mechanism for promoter recognition. However, upon reviewing studies that highlighted the τA subcomplex's enhanced structural stability in yeast [37], I decided to also include the human τA subcomplex in my research. This approach was taken to gain a detailed understanding of both the τA subcomplex and the complete hTFIIIC complex in the context of transcription initiation.

First, my project involved acquiring the human TFIIIC genes from the DNA repository DNASU. I then proceeded with the cloning of individual human TFIIIC subunits and their assembly into the hTFIIIC complex and τA subcomplex, utilizing a baculovirus expression vector. This section also details the protocol developed for purifying these complexes. A pivotal component of the project was optimizing conditions for cryo-electron microscopy (cryo-EM) analysis. Once the samples were stable for cryo-EM, I continue with the acquisition of high-resolution structural data to try to solve the molecular architecture of TFIIIC and its τA subcomplex, with a focus on their interactions with promoter regions.

2.1. Expression and purification of recombinant human τA subcomplex

2.1.1 Small scale purification of τA subcomplex
Before conducting any pulldown assays, it's important to mention that Helga Grötsch was responsible for preparing the τA subunits for expression in insect cells (see table 4.1 for gene and subunit names). Her expertise facilitated the insertion of the Twin-Strep-tag at the N-terminus of GTF3C3, followed with the assembly of the τA subunits into the pBig1 vector using the biGBac system method, a process detailed in section 4.5.2. At the same time, H.G. continued with the project, which involved preparing the bacmid DNA and performing the first transfection to produce the V_0 virus. This is described in more detail in section 4.6.1. These initial steps were crucial for the later stages of the project, which I then took over.

After successfully generating and titrating the V_1 virus, I focused on finding the best insect cell lines for large-scale expression, aiming for a 6-liter culture. This part of the project included transfecting 25 ml cultures of SF21 cells at 0.5×10^6 cells/ml and High Five (Hi5) cells at 1×10^6 cells/ml with a 1:1000 ratio.

After harvesting the cells, I explored two different affinity purification methods to increase the yields of the τA subcomplex. The first method, which is similar to the one used in reference [37], involved using Ni-agarose and Strep-Tactin Sepharose™ High Performance beads. This protocol utilized the His-tag on GTF3C5's N-terminus and the Twin-strep-tag on GTF3C3's N-terminus. The details of the τA subcomplex

construct are explained in section 4.5.2. The second method employed only Strep-Tactin beads. The aim was to determine whether reducing the number of purification steps could improve efficiency and increase yield.

The results of these tests are shown in figure 2.1, highlighting two main findings. First, the method using both Ni-agarose and Strep-Tactin beads produced a purer τA subcomplex, but yielded less product compared to using only Strep-Tactin beads. Second, when using only Strep-Tactin beads, the High Five cells were more effective than SF21 cells in producing τA. These results give important information about the differences in purification methods and which cell lines work best for producing τA.
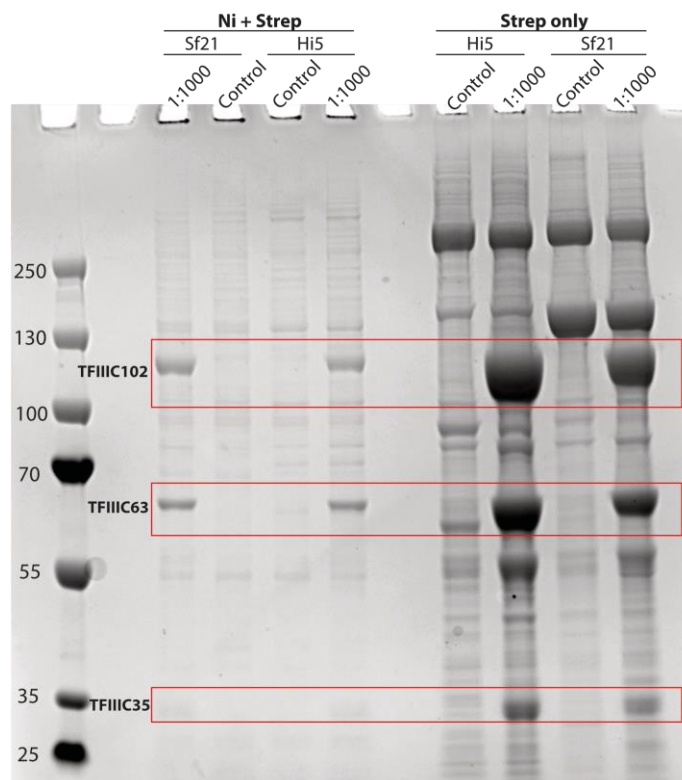
Figure 2.1: Comparative pulldown efficiency of human τA subcomplex in High Five (Hi5) and SF21 cells with 1:1000 virus titration, τA subunits marked by red rectangle; control lanes without virus included for comparison.

2.1.2 Big scale purification of τA subcomplex

After seeing the results from previous experiments, I decided to scale up to a 6L cell culture using High Five cells for a larger expression of the τA subcomplex. The pulldown assays showed good yields of the τA subcomplex, but there was still a need to reduce impurities in the sample. To do this, after the strep tag elution step, I used an ionic exchange column, specifically a Capto HiRes Q 5/50 column (see figure 2.2 A). This step was crucial for increasing the homogeneity of the sample. After this, the τA subcomplex was concentrated and then analyzed using SDS-PAGE. The purification yielded a total of 3.9 mg of protein from 6 liters of culture medium. The sample was also sent for mass spectrometry to identify the bands, as shown in figure 2.2 B.
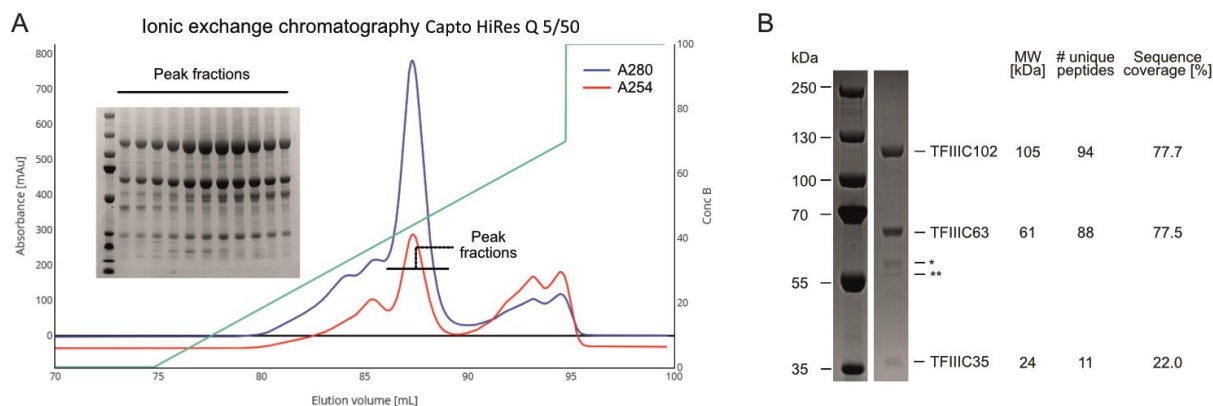
Figure 2.2: Purification of τA subcomplex. A. Ionic exchange chromatogram displaying various peaks, with corresponding SDS-PAGE for pooled fractions containing τA subcomplex. **B.** SDS-PAGE of concentrated sample from the ionic exchange step stained with Coomassie blue, alongside mass spectrometry results.

2.2 Optimization of buffers for cryo-EM studies of τA subcomplex

Following the purification of a homogeneous τA subcomplex sample, the subsequent step involved its stabilization for cryo-EM. This process required screening a variety of grids and detergents, as detailed in section 4.8.1, utilizing the Talos™ Arctica™ microscope (ThermoFisher). Illustrating this, Figure 2.3 displays three micrographs collected under different protein concentrations, using two types of grids, with the same concentration of the detergent Octyl-beta-Glucoside (OG).
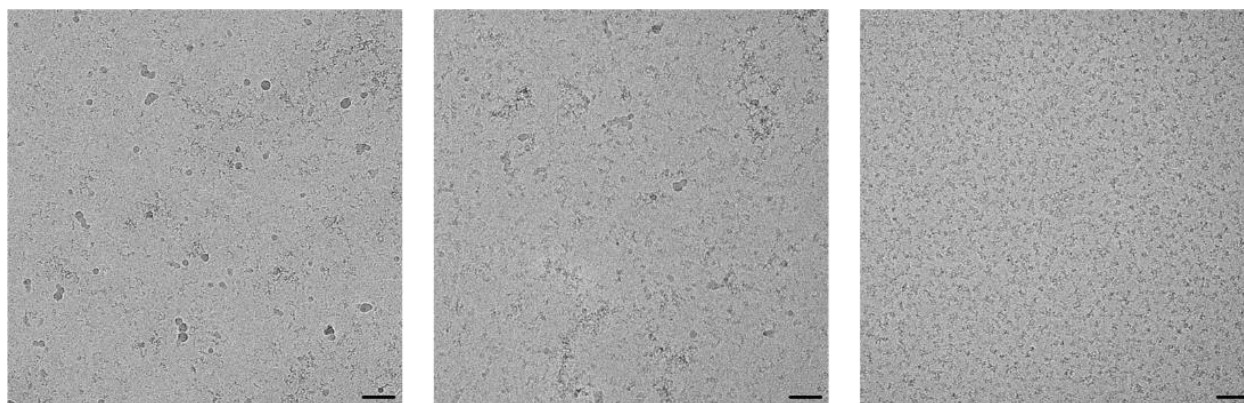
Figure 2.3: Stabilization conditions for τA subcomplex in cryo-EM. Left: τA subcomplex at 0.6 mg/ml in 0.1% OG, using Quantifoil R2/1 Cu 200 grids. Middle: τA Subcomplex at 1.2 mg/ml in 0.1% OG, on Quantifoil R2/1 Cu 200 grids. Right: τA Subcomplex at 0.48 mg/ml in 0.1% OG, using Quantifoil R1/2 Au 400 grids. Scale bar represents 500 Å. Micrographs collected using a 92,000x magnification, corresponding to a pixel size of 1.566 Å.

On the process of preparing homogeneous samples, various parameters such as detergent type, grid selection, and protein concentration are critical to evaluate. When considering the variables from the conditions presented in Figure 2.3, three key factors emerge: grid type, hole size, and protein concentration. Previous research has demonstrated that for a 750 kDa asymmetric multi-protein complex, grids with larger hole sizes of 1.2 and 2 µm led to disassociation, unlike smaller holes in lacey grids [106]. This observation is mirrored in my study, where grids with smaller holes (1 µm) proved more

effective. Additionally, the concentration of protein applied to the grid is a determinant of stability even before application to the grid [107], a finding that might be also reflected in my research. Therefore, it's advisable to rigorously test various conditions for new protein complexes or samples.

2.3 Cryo-EM studies of τA subcomplex

After identifying the appropriate conditions and utilizing the same grid as in the buffer conditions test (see Figure 2.3, right micrograph), a dataset of 1,344 micrographs was collected at a magnification of 120,000x, corresponding to a pixel size of 1.237 Å. The 2D classification revealed promising results showing the potential τA subcomplex in multiple orientations (Figure 2.4 A). Subsequently, I used our in-house Titan Krios 2 for further data acquisition, aiming for a higher resolution structure. The processed data revealed 2D classes with high-resolution features (Figure 2.4 B), showing 'fuzzy' regions that suggest flexibility in the human τA complex. Using ab initio reconstruction based on these 2D classes, followed by non-uniform refinement in cryoSPARC, I generated a 3D model with a resolution of 6.8 Å (Figure 2.4B). Attempts to accurately fit the individual AlphaFold2-predicted τA subunits into this map were unsuccessful due to the current resolution. This difficulty in fitting these predicted structures into the density map may be attributed to potential conformational changes in the subunits upon interaction, which are different from their isolated forms. Additionally, the low-resolution cryo-EM map obtained was unexpected, especially compared to the well-defined yeast τA subcomplex [37]. Despite efforts to improve the map resolution, including particle training/picking with TOPAZ and different classification strategies for obtaining a more homogeneous τA particle population, the inherent flexibility, small size (180 kDa) in its monomeric form, and lower symmetry likely contributed to the low resolution of the τA Subcomplex. Similar factors were implicated to the low resolution in the cryo-EM studies of the cancer target IDH1 (isocitrate dehydrogenase) [108].
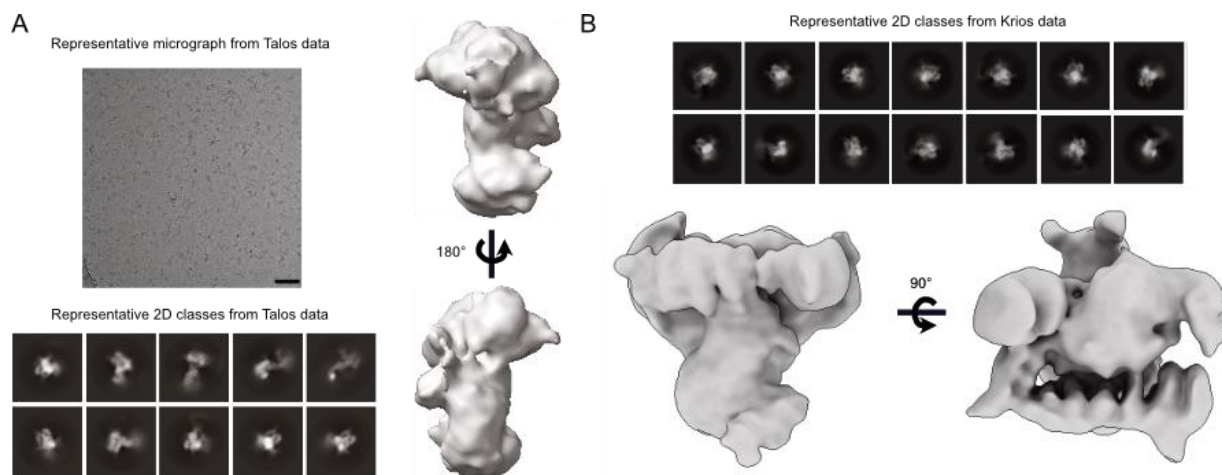
Figure 2.4: Cryo-EM Studies of τA subcomplex. A. A representative micrograph collected using the Talos™ Arctica™ microscope is shown, along with selected 2D class averages from initial data processing. To the right, the reconstructed 3D volume from chosen particles is displayed. Scale bar: 500 Å. **B.** Displayed are 2D class averages from the Krios dataset, illustrating high-resolution features. Protein flexibility is evident, indicated by fuzzy densities around well-defined protein regions. The 3D volume, derived from further data processing, achieved a resolution of 6.8 Å.

2.4 Cloning and purification of human TFIIIC

While working with the cryo-EM data of the τA subcomplex, I faced difficulties in obtaining a high-resolution map. At the same time, I successfully cloned the human τA- and τB genes using the biGBac system [109] (Figure 2.5 A). The decision to replace the commonly used Twin-Strep-tag at the C-terminus of TFIIIC's largest subunit, TFIIIC220 (τ138 in yeast), was influenced by a recent study that used a 3xFlag-tag for the hTFIIIC cloning strategy [93]. After assembling the pbiG2 plasmid with all TFIIIC genes, I continue me with producing a recombinant baculovirus. Initially, to test the expression of the hTFIIIC complex, I used High Five cells. Surprisingly, the complex appeared very clean when only anti-Flag beads were used in the pulldown assay (Figure 2.5 B). These results led me to scale up the expression using the same type of insect cells.
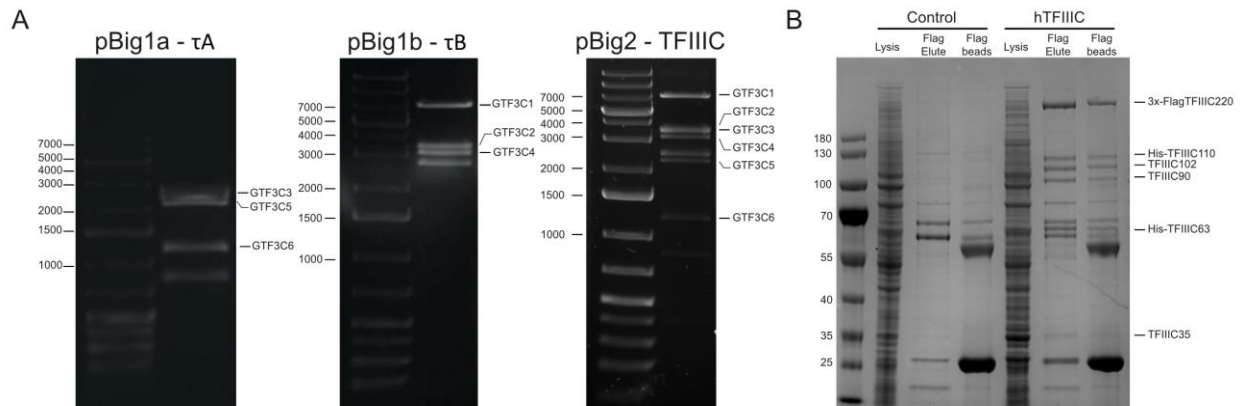
Figure 2.5: hTFIIIC gene cloning and expression in insect cells via baculovirus. A. Analysis of the pBig1 and pBig2 vectors containing all the hTFIIIC genes by restriction digest (SwaI) and gel electrophoresis. The left panel depicts τA genes integrated into the pBig1a plasmid via the biGBac system. The middle panel illustrates τB genes assembled into the pBig1b plasmid. The right panel displays all hTFIIIC genes successfully incorporated into the pBig2 plasmid. **B.** An SDS-PAGE stained with Coomassie blue used after a small-scale pulldown assay using High Five (Hi5) cells is shown. A control sample is included for comparison. The subunits and their respective tags are labeled on the right.

Although small-scale purification (pulldown) using 25 ml of cells at a density of 10^6 cells/ml indicated promising expression results, scaling up to 6 liters presented several challenges. Firstly, the protein yield per expression varied between 80 - 120 µg, markedly less than the nearly 4 mg obtained for the human τA subcomplex as described previously. Secondly, persistent DNA contamination was observed during protein purification. This contamination was difficult to eliminate due to the overlap of the protein contaminant or DNA with the elution volume of hTFIIIC, as evidenced by the 260/280 absorbance ratio detailed in the table in Figure 2.6 A. Implementing an additional size exclusion chromatography step was not feasible due to the low total protein yield at this stage. To address these issues, one of the strategies was to switch the expression host to SF21 cells. As shown in Figure 2.6 B, this change effectively resolved the DNA contamination issue, with the absorbance ratio of the concentrated sample after the Capto HiRes Q 5/50 purification step nearing 0.6, indicating a sample largely free of DNA contamination.
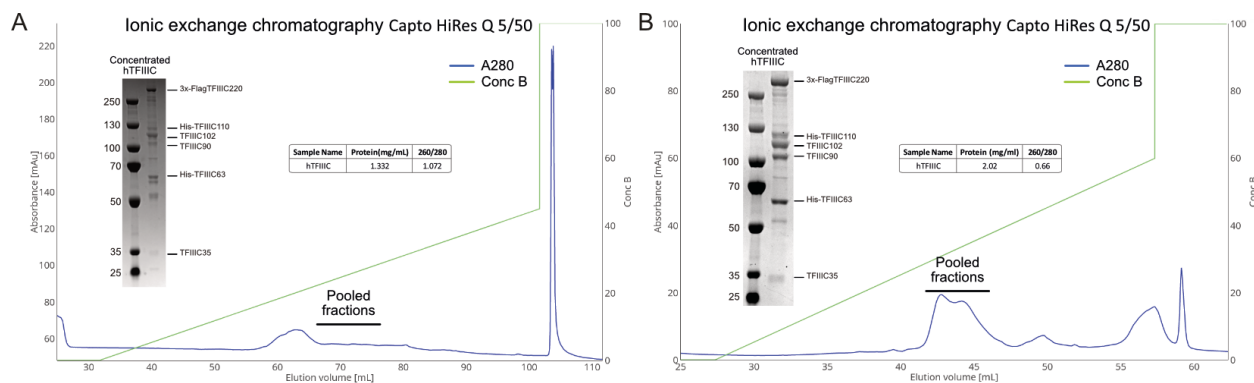
Figure 2.6: Comparative ionic exchange chromatography of hTFIIIC from High Five and SF21 Cells. A. Ionic exchange chromatogram of the sample derived from High Five cells. This panel includes the pooled fractions used for hTFIIIC concentration and a table displaying protein concentration alongside its 260/280 absorbance ratio. **B.** Ionic exchange chromatography results for the hTFIIIC sample extracted from SF21 cells. It shows the pooled fractions for hTFIIIC concentration and a table with protein concentration and 260/280 absorbance ratio.

It is important to note that following each purification step, the fractions containing hTFIIIC were pooled and concentrated. Subsequently, a buffer exchange was performed to remove glycerol and reduce the high salt concentration, which could potentially compromise the sample's stability during cryo-EM sample preparation. The sample was then immediately used for cryo-EM preparation. Any remaining sample was flash-frozen and stored at -80°C for future biochemical or biophysical experiments. Furthermore, the purified and concentrated sample was subjected to mass spectrometry to identify all subunits and impurities, as shown in Figure 2.7 A.

2.5 Preparing hTFIIIC for cryo-EM studies

To assess the structural integrity of the purified hTFIIIC and investigate its binding to the TRR-TCT3-2 gene, mass photometry (MP), a technique for detecting individual unlabeled molecules through light scattering, was utilized [110]. More details about the MP experiment can be found in section 4.4. The calculated theoretical molecular weight (MW) of the intact hTFIIIC complex is approximately 624 kDa. This matches with the experimentally determined MW of 654 ± 52 kDa by MP (refer to Figure 2.7 B). In a parallel experiment adding the specific tRNA gene, mass photometry revealed two overlapping peaks: one for hTFIIIC alone (660 ± 29.4 kDa) and one for the hTFIIIC-DNA complex (701 ± 46 kDa), as depicted in Figure 2.7 C. This latter peak coincides with the sum of the theoretical MW of hTFIIIC (624 kDa) and the DNA (expected MW: 64 kDa). Motivated by these results, I advanced to further investigate the DNA-binding properties of hTFIIIC using two additional assays. An Electrophoretic Mobility Shift Assay (EMSA) demonstrated a concentration-dependent shift in DNA migration in the presence of hTFIIIC, with minimal binding observed at a 0.5:1 protein:DNA ratio (Figure 2.7 D). The DNA concentration in this assay (0.2 μM) exceeds the estimated K_D, rendering these conditions suboptimal for K_D determination (determined with filter binding assay below).

Our filter binding assay, performed by Florence Baudin, further confirmed the high affinity of hTFIIIC for the TRR-TCT3-2 gene. The dissociation constant (K_D) for the interaction was determined to be 40 nM. This finding is similar to previous results observed in yeast TFIIIC binding to a specific tRNA(tyr), SUP4,

where the equilibrium binding constant was determined at 1.0×10^{10} M^{-1} (equivalent to 0.1 nM affinity) [35].
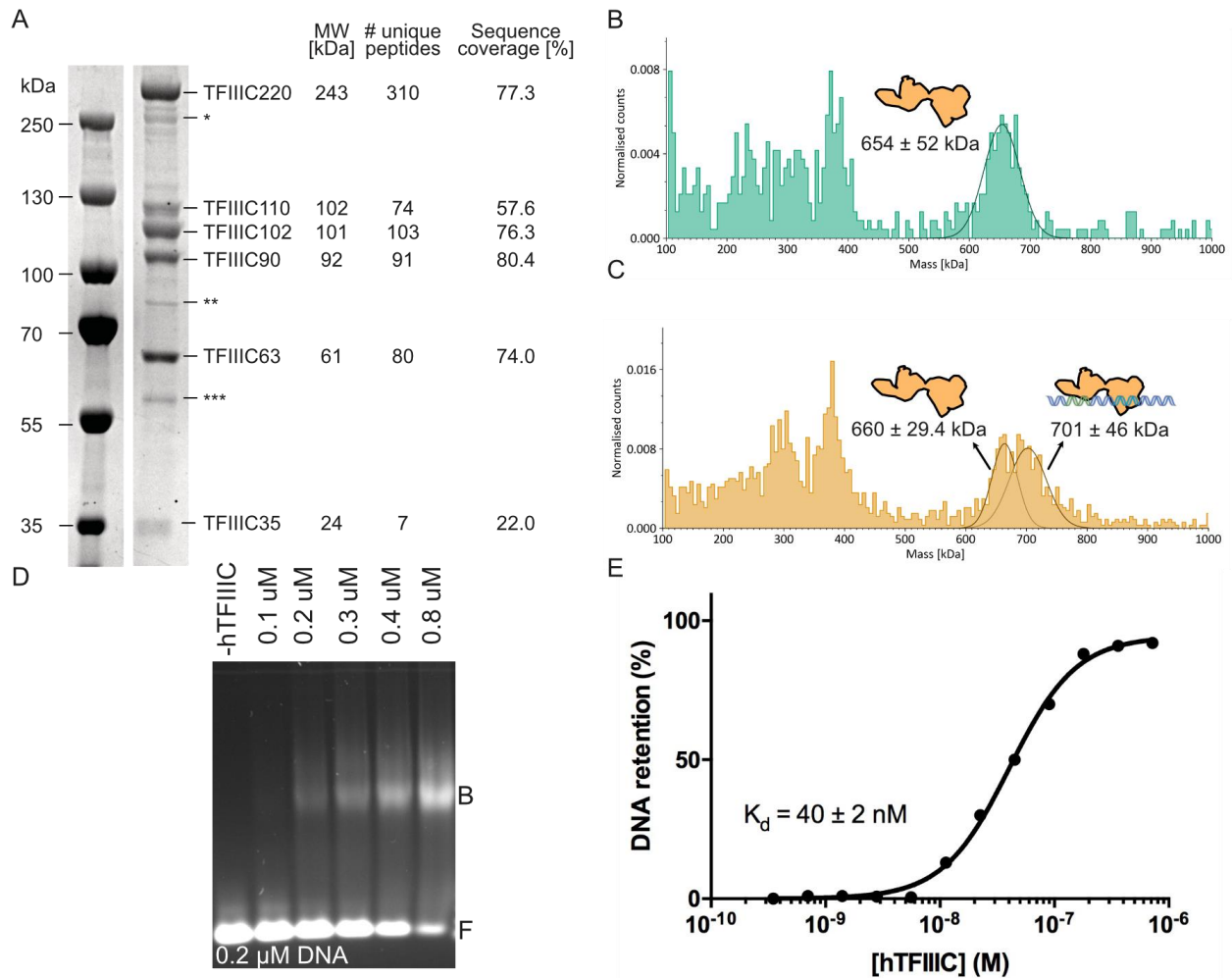
Figure 2.7. Identification, complex integrity of hTFIIIC and testing hTFIIIC-DNA interaction. **A.** SDS-PAGE of the purified hTFIIIC complex. All the subunits were identified by mass spectrometry. (*) indicates TFIIIC220 degradation products, (**) shows the presence of 70kDa heat shock protein from the expression host, and (***) shows alpha-beta tubulins from the host. **B.** Mass photometry analysis of hTFIIIC. The mass distribution histogram is displayed, alongside a schematic of hTFIIIC corresponding to the peak with a molecular weight of 654 ± 52 kDa. **C.** Mass photometry analysis of DNA-bound hTFIIIC. The histogram shows mass distribution for the hTFIIIC-DNA sample. Schematics of hTFIIIC unbound (left) and bound (right) to DNA are presented next to the peaks corresponding to 660 ± 29.4 kDa and 701 ± 46 kDa, respectively. **D.** EMSA confirming the interaction between hTFIIIC and the TRR-TCT3-2 gene. 'F' indicates free DNA and 'B' indicates DNA bound to hTFIIIC. **E.** Filter binding assay determining the dissociation constant (K_D) between hTFIIIC and the TRR-TCT3-2 gene, with binding data fitted using a Hill equation. The estimated dissociation constant (K_D) for the interaction was determined to be 40 nM, with a standard error of 2 nM. Panels A to E were slightly modified from the reference [66].

Upon confirming that hTFIIIC formed a complete and active complex, as evidenced by its binding to the tRNA gene, I proceeded to test various buffer conditions for cryo-EM sample preparation, similar to the approach detailed in section 2.2 for the τA subcomplex. A series of micrographs (Figure 2.8 A and B) demonstrate the impact of different NaCl and KCl concentrations and the addition of glutaraldehyde (GA) on complex stabilization. For the hTFIIIC without DNA sample, salt concentrations below 200 mM NaCl led to increased aggregation, likely due to a "salting in" effect. This effect, where moderate salt concentrations increase protein solubility, has been previously observed in the *E. coli* complex I, which showed significant aggregation at 50 mM NaCl, greatly reduced at 250 mM NaCl [107]. Additionally, my efforts to use GA, previously used to stabilize large complexes for cryo-EM, were unsuccessful in the case of hTFIIIC [111], [112].

Furthermore, my initial strategy to stabilize hTFIIIC with DNA involved lowering the salt concentration. Preliminary findings in yeast, based on DNAse I protection experiments, suggested an optimal salt concentration of 135 mM KCl for the interaction between TFIIIC and tRNA$_3^{Glu}$. At concentrations above 200 mM, a lack of footprinting indicated complex dissociation [113]. Similarly, a comparison between wild type yeast TFIIIC and a mutant (using comparable volumes of cell extracts) revealed an optimal binding affinity to tRNA$_3^{Glu}$ at 150 mM KCl in gel shift assays[114]. Additionally, a study comparing active and inactive forms of human TFIIIC at various salt concentrations using gel mobility shift assays revealed a lower "optimal salt concentration" for human TFIIIC of 70 mM, compared to yeast [71]. However, these optimal conditions were not suitable for the human TFIIIC complex in cryo-EM studies.
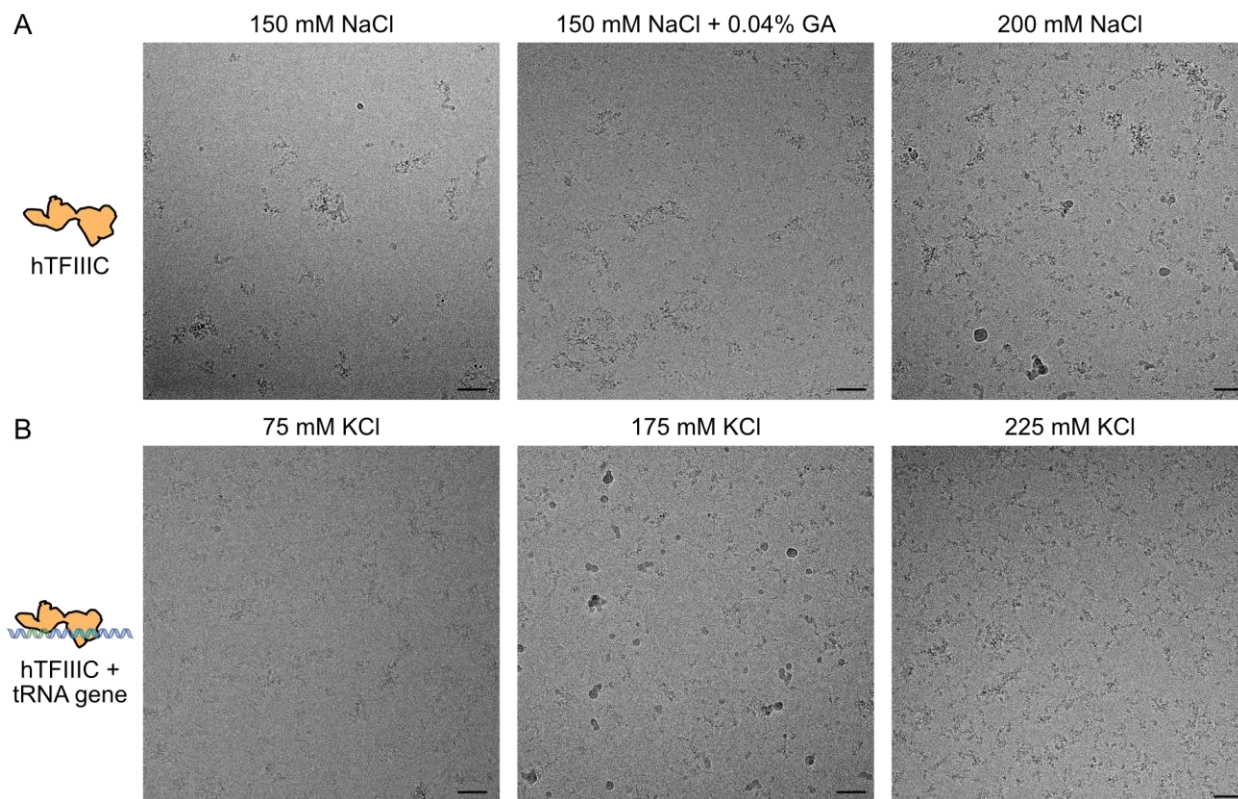
Figure 2.8: Buffer condition optimization for hTFIIIC and hTFIIIC-DNA complex. A. Analysis of hTFIIIC alone (left micrograph) with 150 mM NaCl shows aggregation on an Ultrafoil R2/2 Au200 grid; addition of 0.04% glutaraldehyde under similar conditions results in comparable aggregation on a Quantifoil R 2/1 + 2nm Cu200 grid (middle micrograph). Increasing NaCl concentration to 200 mM, however, leads to decreased aggregation and more visible individual particles (right micrograph). **B.** Evaluation of KCl concentration on hTFIIIC-DNA complex stability: At 75 mM KCl, aggregation is observed on an Ultrafoil R2/2 Au200 grid (left micrograph); increasing KCl concentration to 175 mM results in both aggregation and free DNA (middle micrograph), whereas at 225 mM KCl, there is a marked improvement in stability, evident by reduced aggregation and free DNA (right panel). Detailed sample preparation is outlined in section 4.9.1, except for any variations specified in this figure. Salt concentrations and additives are labeled on each micrograph. Scale bar indicates 500 Ångströms. Micrographs were taken at a magnification of 92,000x, corresponding to a pixel size of 1.566 Å.

2.6 Cryo-EM structures of the hTFIIIC and hTFIIIC-DNA complex

After establishing appropriate conditions to stabilize hTFIIIC for cryo-EM, I collected 11,088 movies using our Titan Krios microscope equipped with a K3 camera; refer to Table A1 to see more details about data acquisition. The workflow for generating high-resolution cryo-EM maps of the τA and τB subcomplexes is detailed in Figure 2.9 A. Initially, WARP-selected particles were sorted using 2D classification, revealing two distinct particle sets. Subsequent TOPAZ training and picking steps were applied to each population separately, resulting in an increase of 2 to 2.5 times in particle count compared to the initial amount of particles used. To further classify the particles, for the τA subcomplex, multiclass ab initio reconstruction was performed, while the τB subcomplex underwent both ab initio and heterogeneous refinement. This methodology achieved resolutions of 3.8 Å for the τA subcomplex and 3.4 Å for the τB subcomplex

(Figure 2.9 B). Due to issues with preferred orientation, modeling of the τA subcomplex was not achievable. However, for the τB dimeric complex, modeling was successfully performed using the Alphafold2 predicted structures of its individual subunits as starting points (see figure 2.9 C). Due to the high local resolution in the core region, which reached up to 3.2 Å, as detailed in Figure 2.9 D, the model building was successfully achieved. Comparative analysis of this model with the hTFIIIC-DNA complex will be described in the following section.

To solve the high-resolution structures of the hTFIIIC-DNA complex, I performed a workflow similar to that used for the "hTFIIIC without DNA" sample. I collected 11,025 movies with the Titan Krios microscope, with more details on data acquisition in section 4.9.2. The initial processing steps of this data revealed two distinct particle groups, likely corresponding to the τA and τB-dimer subcomplexes, as observed in the previous sample without DNA. However, the 2D classification indicated these classes differed slightly from the previous sample. As shown in Figure 2.10 A, processing the data led to resolutions of 3.5 Å for τA and 3.2 Å for τB-dimer maps. FSC curves are illustrated in Figure 2.10 B. Contrary to the hTFIIIC sample without DNA, the τA subcomplex showed no preferred orientation, enabling atomic model construction. The use of cryoDRGN significantly improved the sorting of τB-dimer particles (performed by Luis Hauptmann), enhancing the resolution at the winged helix domains, which interact with the DNA. This was a key step in understanding the structural mechanism of promoter recognition by hTFIIIC. Despite applying the same software to τA particles, it did not result in similar improvements. Additionally, the high local resolutions obtained for both subcomplexes aided in constructing their high-resolution structures, as depicted in Figures 2.10 C and D.
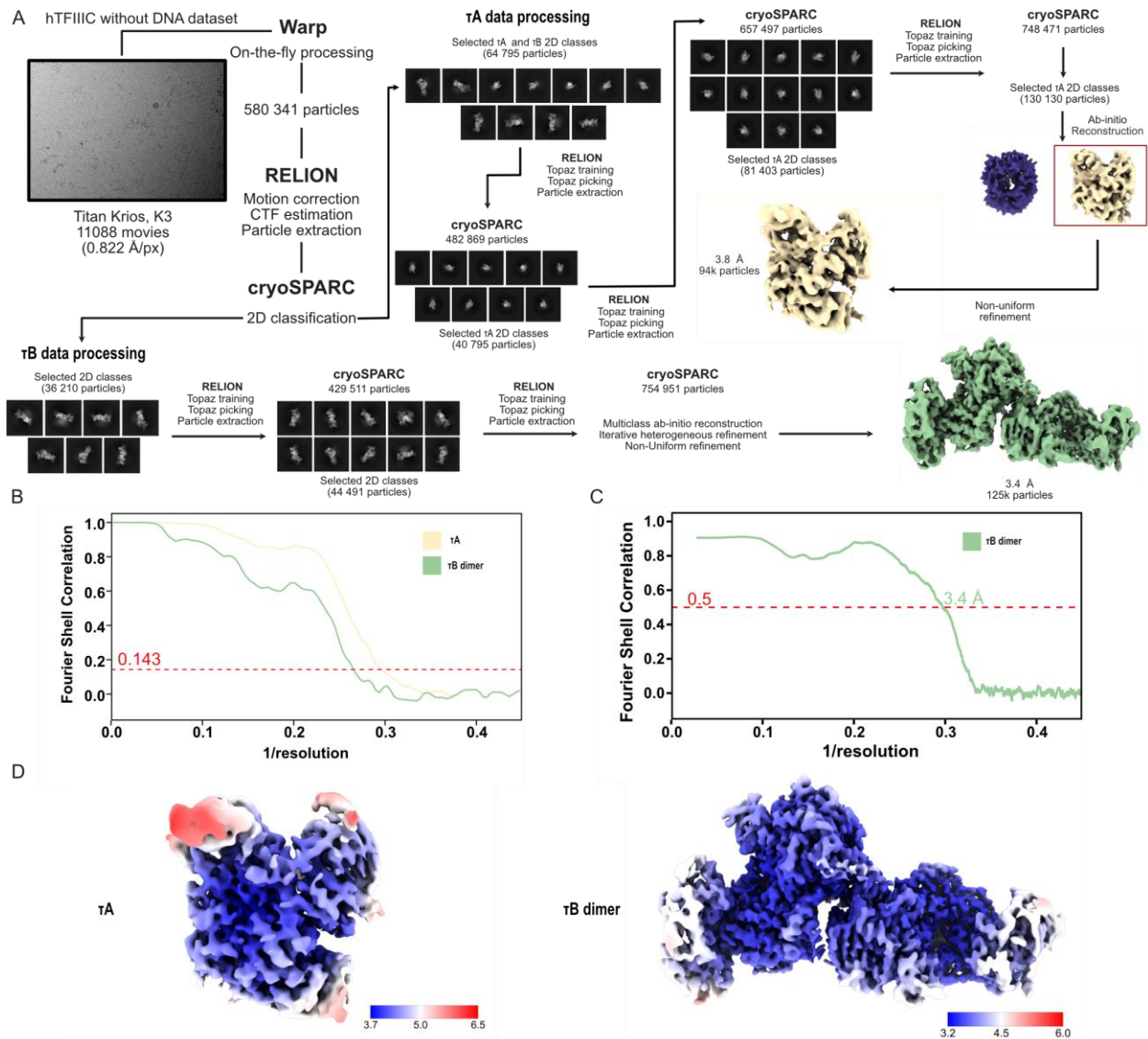
Figure 2.9. Cryo-EM workflow for data processing and assessment of hTFIIIC quality. **A.** The processing pipeline for the "hTFIIIC without DNA sample" obtained from the Titan Krios dataset is detailed. This process includes a representative micrograph, 2D class averages, and 3D classifications from ab-initio reconstruction, along with non-uniform refinement maps of τA (top right) and τB-DNA (right) as produced in CryoSPARC. **B.** Fourier Shell Correlation (FSC) curves of the τA and the dimeric τB map yield final resolutions of 3.8 Å and 3.4 Å, respectively, obtained at an FSC threshold of 0.143. **C.** An FSC was performed between the dimeric τB map and its corresponding model to evaluate the map-to-model fit, with an FSC criterion of 0.5. **D.** The local resolution for both τA and the dimeric τB map was calculated using the RELION software. Panel A to D was adapted from reference [66].
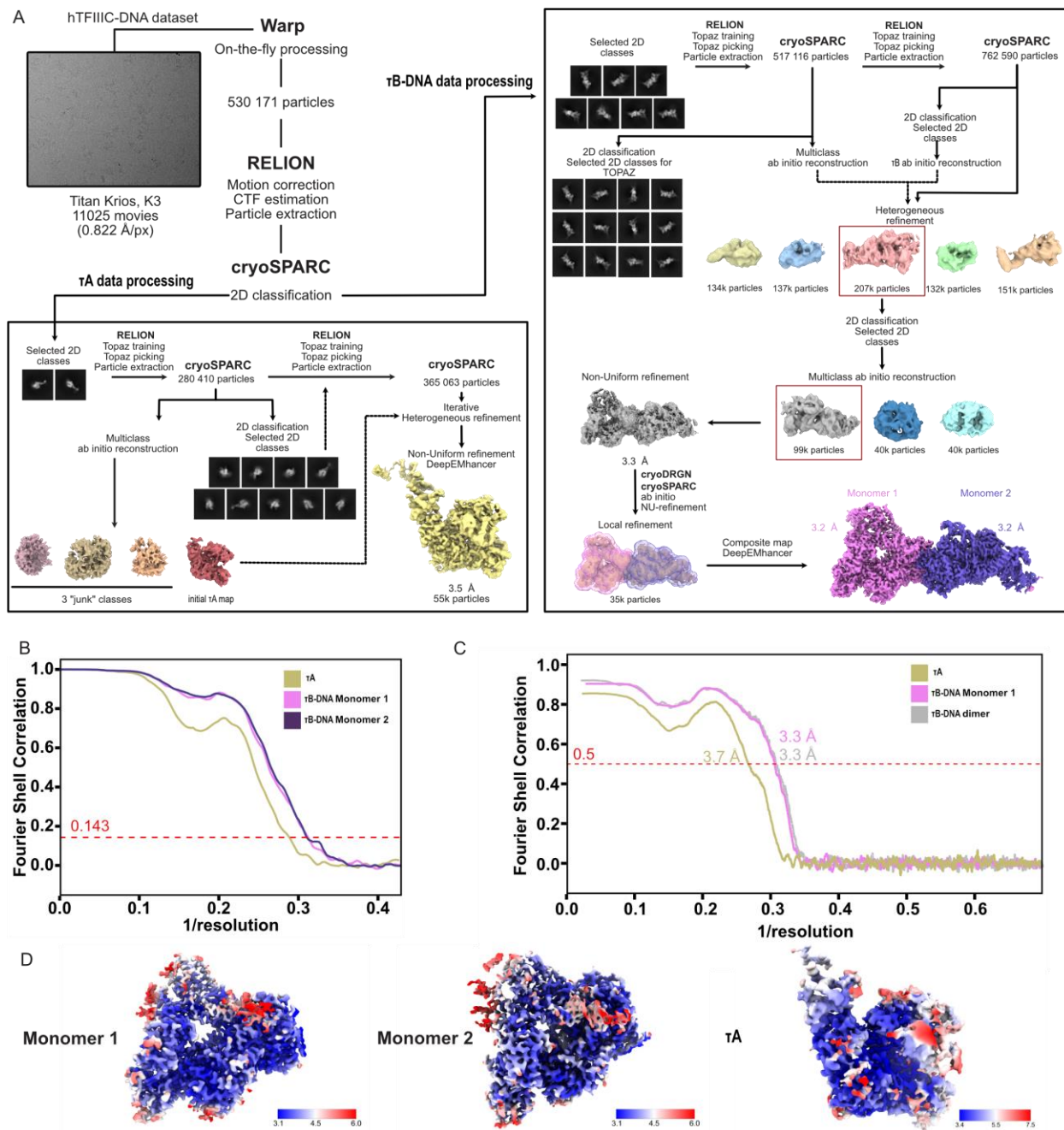
Figure 2.10. Processing of cryo-EM data and quality evaluation for hTFIIIC-DNA complex. A. Processing steps for the hTFIIIC-DNA dataset from Titan Krios, including a representative micrograph, 2D class averages, 3D classifications from heterogeneous refinement, and non-uniform refinement maps of τA (bottom left) and τB-DNA (right) with post-processed DeepEMhancer maps in CryoSPARC. **B.** FSC curves of τA, monomer 1, and monomer 2 of the dimeric τB-DNA map show final resolutions of 3.5 Å, 3.2 Å, and 3.2 Å, respectively, using an FSC criterion of 0.143. **C.** FSC was computed for map-to-model fit assessment, with an FSC threshold of 0.5. **D.** Estimation of local resolution for τA, monomer 1, and monomer 2, performed with RELION. Panel A to D was adapted from reference [66].

2.6.1 Overall structures of human τA and τB subcomplexes

Following building of the model, the domain organization within the human TFIIIC subunits—TFIIIC220, TFIIIC110, and TFIIIC90, which constitute the τB subcomplex, and TFIIIC102, TFIIIC63, and TFIIIC35, forming the τA subcomplex—is depicted in Figure 2.11 A.
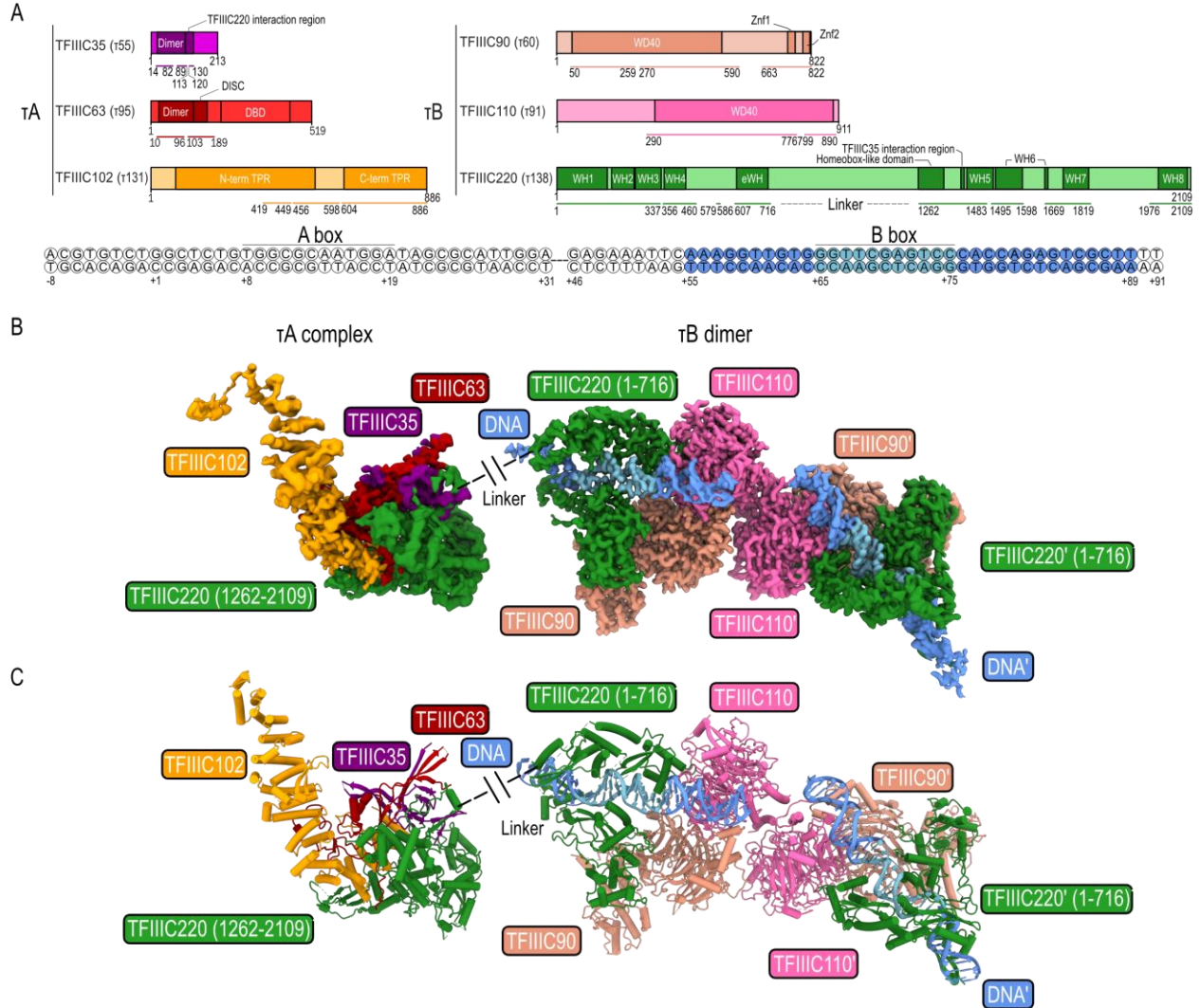
Figure 2.11. Cryo-EM maps and structures of hTFIIIC. A. Domain architecture of hTFIIIC subunits at the top and the tRNA gene utilized in the cryo-EM analysis at the bottom. The nucleotides are numbered from the start of the predicted mature tRNA. Yeast homologs subunits indicated in parentheses. The diagram includes colored bars representing the regions that were built and colored circles denoting the modeled DNA. Key domains are labeled: DBD for DNA binding domain, TPR for tetratricopeptide repeat domain, WD40 for WD40 repeat domain, WH for winged-helix domain, and eWH for extended winged-helix domain. **B.** Cryo-electron microscopy maps of human τA unbound to DNA and the dimeric τB bound to DNA. **C.** Details of the atomic models of human τA unbound to DNA and τB bound to DNA. Panel A to C was adapted from reference [66].

Remarkably, this domain architecture is very similar to that of the yeast TFIIIC's domains, subunits, and τA subcomplex, as reported in earlier studies ([36], [37], [59], [61], [115]), prior to the beginning of my PhD. This similarity is notable given the low sequence identity percentages between the human and corresponding yeast subunits: 6.2% for TFIIIC35 – τ55, 8.1% for TFIIIC220 – τ138, 10.8% for TFIIIC110 – τ91, 11.2% for TFIIIC90 – τ60, 15.7% for TFIIIC63 – τ95, and 19.7% for TFIIIC102 – τ131 [77].

The structural analysis was a collaborative effort involving both myself and Mathias Girbig, who contributed to the publication on hTFIIIC [66]. Upon examining the solved structures, we encountered an unexpected finding: the dimerization of the τB subcomplex. This surprising result is illustrated in Figure 2.11 B and C. Interestingly, this dimerization is also observable in the "hTFIIIC without DNA" sample, as indicated in Figure 2.12. This suggests that dimerization occurs independently of DNA presence and might be influenced by protein concentration, as it was not observed at nanomolar levels in microscale thermophoresis (MP) experiments (refer to Figure 2.7 B and C).
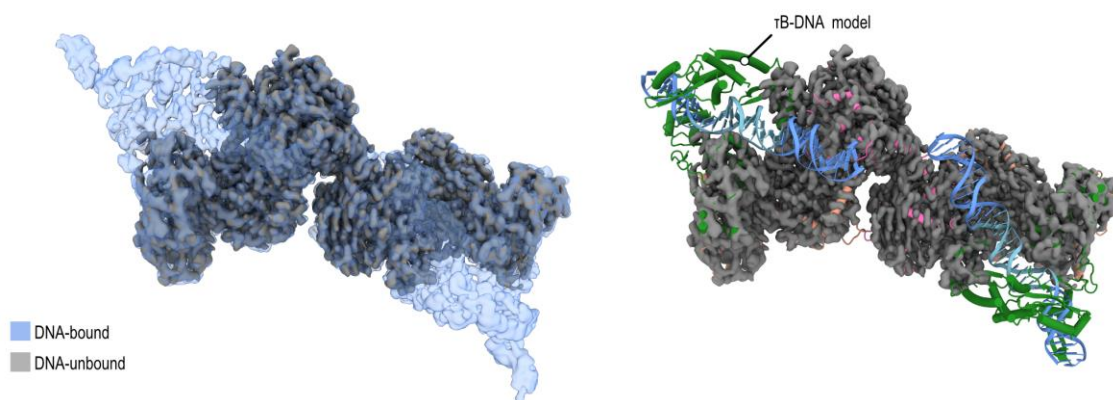
Figure 2.12 Dimerization of τB subcomplex in hTFIIIC with and without DNA. On the left, the cryo-EM density map of hTFIIIC (grey) from the "hTFIIIC without DNA" sample, overlaid with the blue and transparent DNA-bound TFIIIC map. On the right, the superimposition of the hTFIIIC map with the refined DNA-bound TFIIIC model, displayed as a cartoon. Figure adapted from reference [66].

Moreover, the stabilization of the two τB subcomplexes is facilitated by multiple apolar and polar interactions between two copies of the TFIIIC110 subunit, as detailed in Figure 2.13. Considering TFIIIC's involvement in genome organization [116] and its role in modulating the genome through chromatin looping and histone acetylation [93], the specific function of this dimer may be linked to these functions. At present, my analysis will center on the τB monomer, considering that both monomers are structurally identical and the study of this dimerization is constrained by the small amount of sample available after purification.
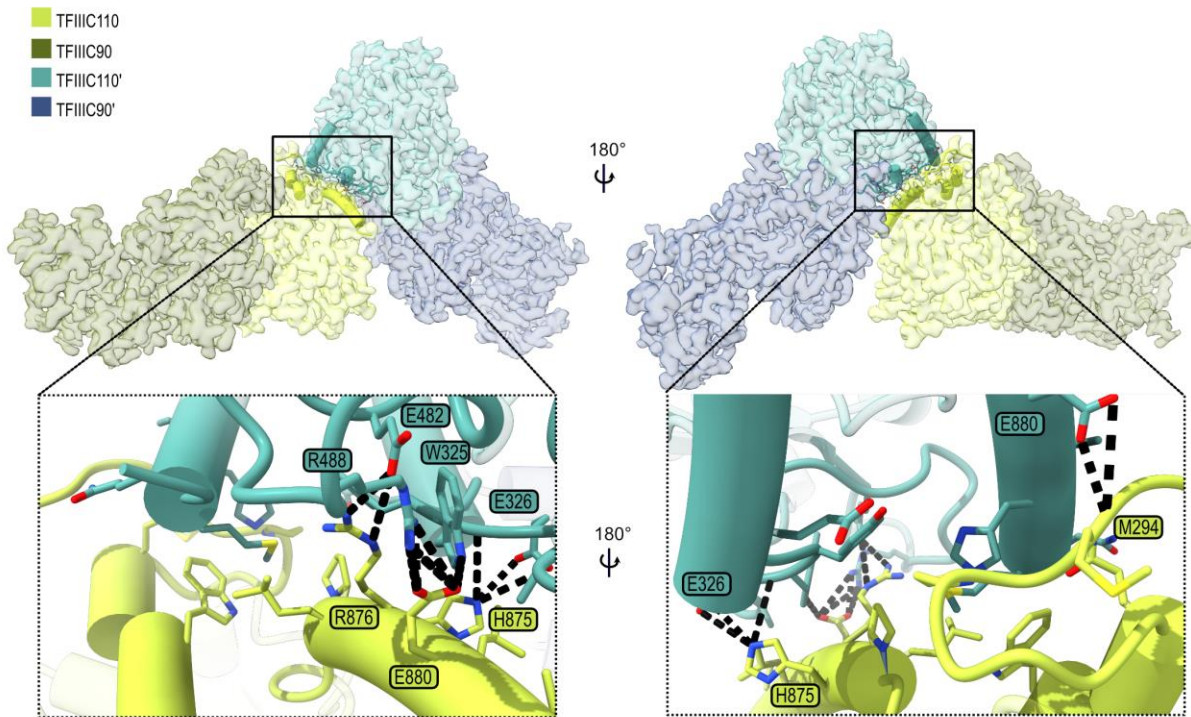
Figure 2.13 τB subcomplex dimer interface. Details of the interface interaction between τB monomers. Insets highlight the amino acids involved in the interaction, with black dashed lines indicating polar interactions and the participating amino acids enclosed in colored boxes. The amino acids contributing to the interactions were identified using ChimeraX's "interfaces" command line and PDBePISA. Figure adapted from reference [66].

2.6.2 Interaction of τB subcomplex with tRNA gene

The core of the human τB subcomplex is constituted by TFIIIC110 and TFIIIC90 (known as τ91 and τ60 in yeast), which form a stable association through their WD40 domains (see figure 2.14 A). A similar interaction is observed in yeast, where the crystal structure of τ60/Δτ91 reveals the β propeller domains of both subunits are aligned perpendicularly to each other [59]. Furthermore, the N-terminal region of the TFIIIC220 subunit (τ138 in yeast) was successfully mapped, comprising roughly one third of the total residues (1 to 716). Containing four winged-helix (WH) domains and one extended WH (eWH) domain, this region forms a crucial part of the human τB structure, as depicted in Figure 2.11 A and Figure 2.15 A and C. This observation contrasts with earlier predictions for the yeast τ138 subunit based on its sequence analysis. Initially, it was hypothesized that the first 68 amino acids from the N-terminal end would constitute a High Mobility Group (HMG) domain, potentially providing this subunit with DNA binding capabilities [52].

The interaction between the WH1 domain of TFIIIC220 (residues 1 to 163) and the WH2 domain (residues 175 to 247) is shown in Figure 2.14 B. Additionally, the WH1 domain of TFIIIC220 binds to two zinc fingers in the C-terminal domain of TFIIIC90, as demonstrated in Figure 2.14 C. This interaction with TFIIIC90 could explain the absence of binding activity in the isolated yeast subunit τ138, suggesting the importance of τ60 for the structural integrity required for its function [52]. At the opposite end of τB's N-terminus, the WD40 domain of TFIIIC110 interacts with the WH4 domain of TFIIIC220, covering residues

367 to 437. In this arrangement, TFIIIC90 and TFIIIC110, together with WH1 domain of TFIIIC220 and the C-terminal zinc fingers of TFIIIC90, form a stable scaffold stabilizing the WH2 and WH4 domains. These domains were ordered even in the absence of DNA, whereas the WH3 and eWH domains were not visible in the cryo-EM map of the DNA-unbound human τB, as illustrated in Figure 2.12.
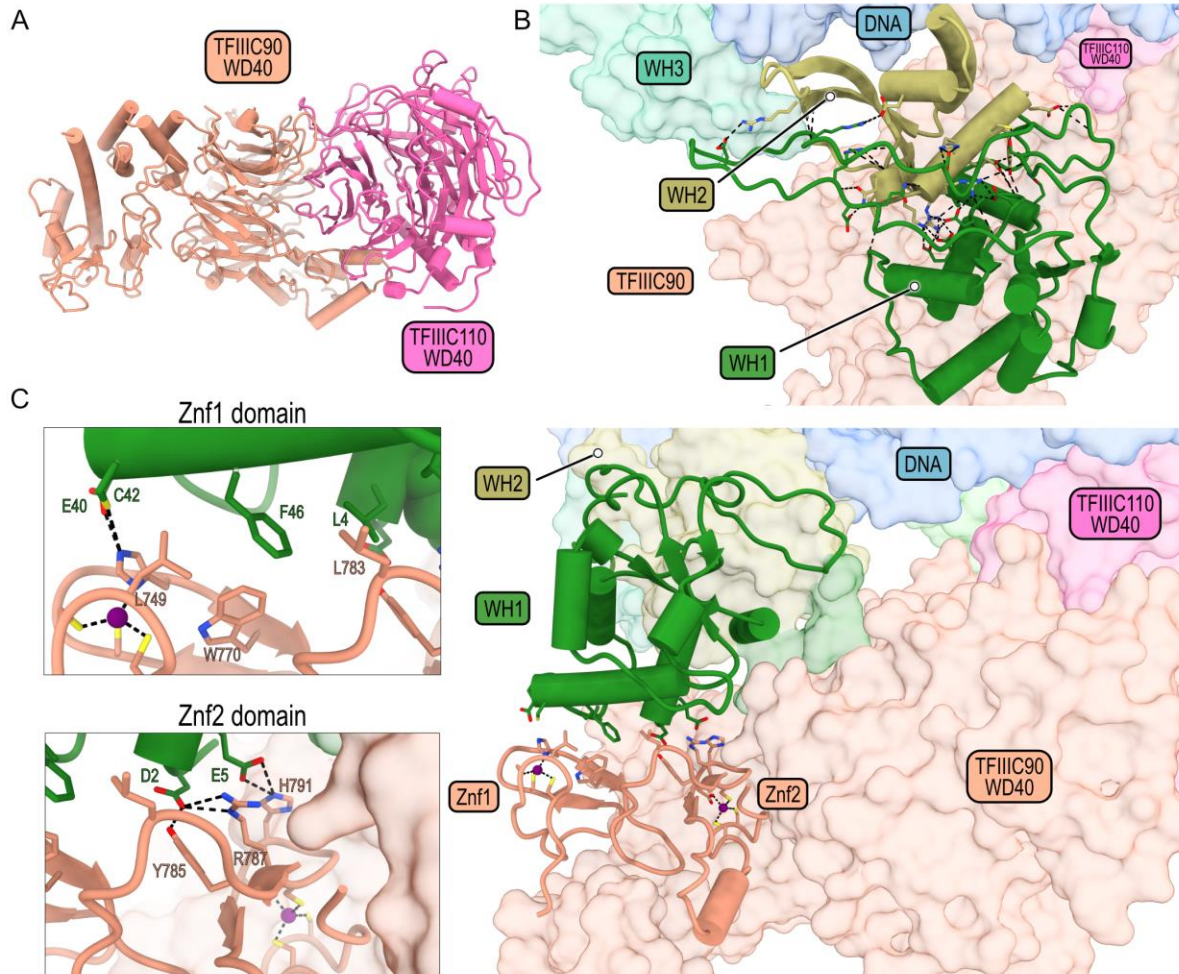
Figure 2.14. Role of τB core in complex stability and DNA binding. **A.** The structural model depicting the τB core, constituted by the interaction between two WD40 domains. **B.** The interaction between the WH1 and WH2 domains, highlighting the side chains of amino acids involved in hydrogen bond formation, with hydrogen bonds depicted as black dashed lines. **C.** The interaction of Zinc fingers (Znf) 1 and 2 (shown in cartoon representation) from the TFIIIC90 subunit with the WH1 domain of the TFIIIC220 subunit. Insets provide a detailed view of the chemical environment facilitating this interaction. Panel A to C was adapted from reference [66].

The elucidation of the τB core structure, focusing on the domains essential for B-box promoter recognition, provides a foundation for a thorough analysis of the structural basis of B-box recognition, as shown in Figure 2.15. The high-resolution cryo-EM map of DNA-bound τB enabled the accurate building of 35 base pairs (55 to 89) of the 3' end of TRR-TCT3-2 tRNA gene sequence, as seen in Figure 2.11 A. The interaction with DNA leads to the formation of the TFIIIC220-WH3 (amino acids 253 to 337) and eWH domains (amino acids 612 to 714), aiding in the recognition of the B-box DNA (nucleotides 65 to 75) in concert with WH2 and WH4, as depicted in Figure 2.15 C. This represents the first instance of a detailed structural description of B-box recognition by the largest subunit of TFIIIC, a significant advancement from prior studies that, using protein-DNA crosslinking experiments, identified this subunit but did not specify individual domains [70], [71].
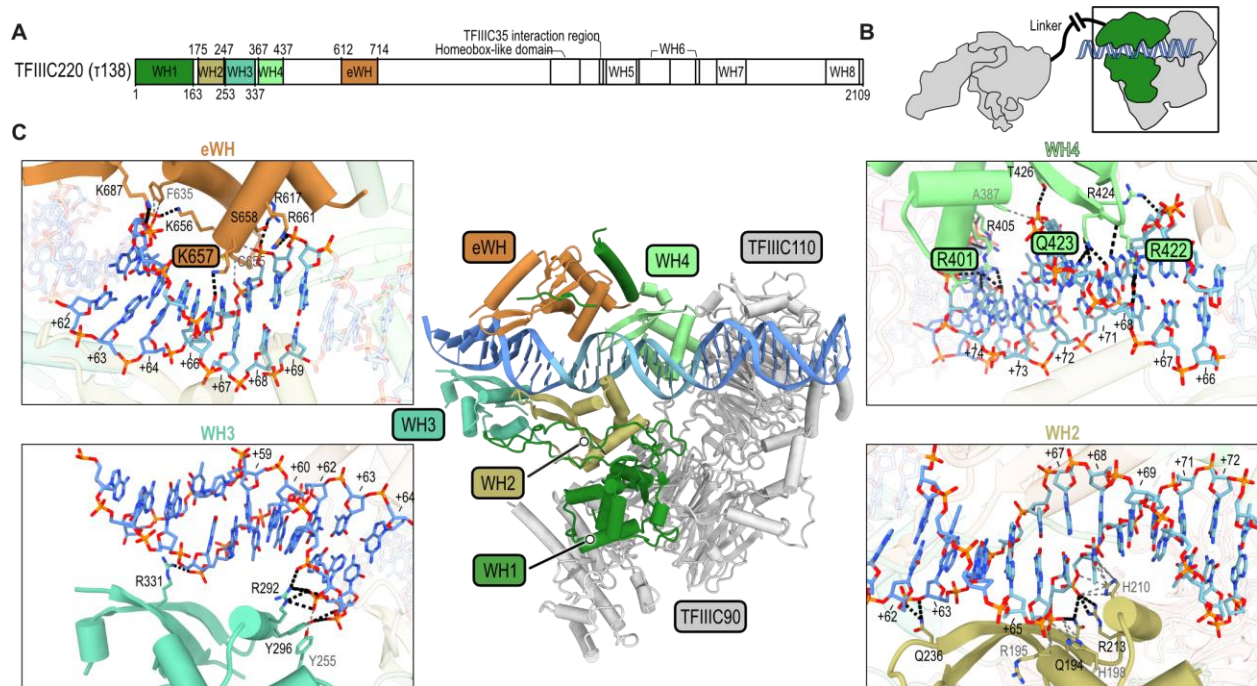
Figure 2.15 . τB subcomplex interaction with B-Box promoter DNA. **A.** Domain architecture of TFIIIC220, with the N-terminal DNA-Binding Domains highlighted in color. **B.** Schematic representation of hTFIIIC bound to DNA, where the N-terminal moiety of TFIIIC220 is depicted in green. **C.** At the center, presents the atomic model of human τB bound to DNA (with the B-box shown in light blue). The interactions between WH2, WH3, WH4, and eWH domains with the DNA are detailed in close-up views. Black dashed lines indicate hydrogen (H) bonds, while gray dashed lines represent apolar contacts. H-bond-forming residues are labeled, and amino acids interacting with the DNA bases are emphasized in colored boxes. Panel A to C was adapted from reference [66].

Further downstream of the B-box, the DNA is additionally stabilized by a positively charged pocket on the TFIIIC110-WD40 domain, as indicated in Figure 2.16. In humans, the absence of this subunit has been associated with a transcriptionally inactive TFIIIC [71], while in yeast, the τ91 subunit (equivalent to human TFIIIC110) has been found to cooperatively bind to DNA with the τ138 subunit [23].
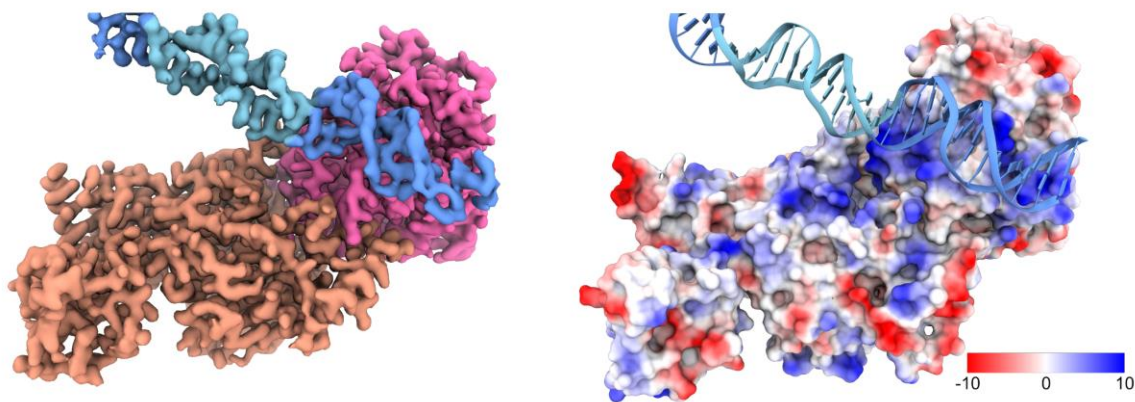
Figure 2.16. TFIIIC110 helps stabilizing DNA binding. On the left, A sharpened map displays the τB scaffold along with the DNA, highlighting the B-box in light blue. On the right, the electrostatic (Coulomb) potential surface representation of the τB scaffold is shown, revealing a positively charged surface on the TFIIIC110-WD40 domain that interacts with the DNA downstream of the B-box. Figure adapted from reference [66].

The WH domains surround the DNA, forming multiple apolar and polar contacts with the bases of the B-box DNA and phosphate backbone (see Figure 2.17 A). Analysis of B-box DNA sequence conservation, performed by Mathias Girbig, reveals that the bases involved in specific contacts (G66, T68, C69, and G70 on the non-template strand; G69, G74, and G75 on the template strand) are conserved not only in human tRNA genes but also across a wide range of eukaryotes (see Figure 2.17 B). This pattern of conservation in tRNA genes has been previously observed in studies focusing on human [90] and mouse [117], corroborating our findings. Additional biochemical DNA binding studies with yeast TFIIIC and the SUP4 gene have demonstrated that mutations C56➔G and G57➔C (corresponding to C69 and G70 on the non-template strand in the human TRR-TCT3-2 tRNA gene) result in a significant decrease in DNA binding affinity [35]. Our structural analysis, combined with these previous experiments, confirms the importance of recognizing these conserved bases. Therefore, TFIIIC recognizes the tRNA gene promoter in an evolutionarily conserved manner, employing a combined DNA shape and sequence readout of the B-box DNA motif.
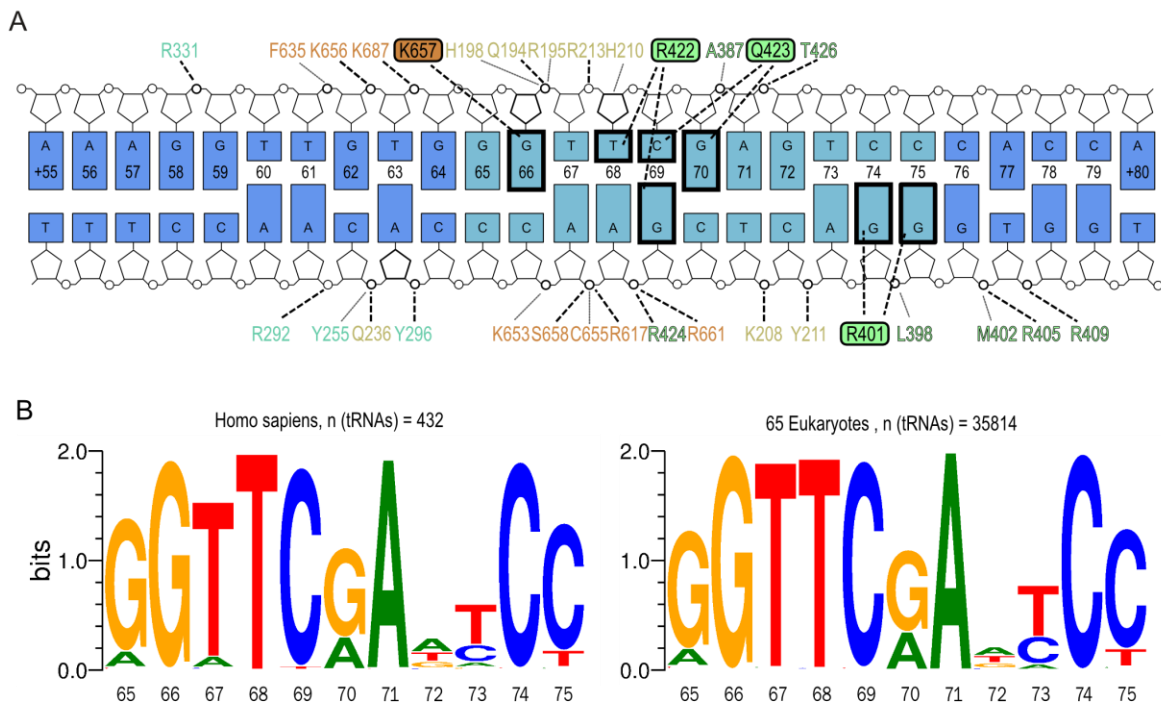
Figure 2.17. TFIIIC220-DNA full interaction and B-box DNA sequence conservation analysis. A. Schematic providing an in-depth view of the protein-DNA interactions observed in the cryo-EM structure. Critical bases that establish hydrogen bonds with specific amino acids are marked in colored boxes, with the color coding for amino acids matching that in Figure 2. 15. Hydrogen bonds are shown as bold, dashed lines, and hydrophobic contacts are represented with thin, grey lines. **B.** Conservation of the B-box DNA sequence within tRNA genes from humans (left panel) and across a broad range of 65 eukaryotic species, including 44 metazoans, 11 fungi, 9 plants, and 1 kinetoplastid (right panel). Panel A to B was adapted from reference [66].

2.6.3 τA and τB are connected by a flexible linker in TFIIIC220

In our cryo-EM analysis, a high-resolution structure of human τA was obtained, which remained unbound to DNA in the presence of the tRNA gene (see figure 2.11 C). This observation is consistent with findings from DNase I footprinting experiments conducted on VA1 genes and $tRNA_1^{MET}$. These experiments showed full protection in the B-box region and minimal interaction at the A-box in the presence of TFIIIC2 (comprising TFIIIC220, TFIIIC110, TFIIIC102, TFIIIC90, and TFIIIC63). The protection in both promoter regions was markedly enhanced with the addition of TFIIIC1, which includes at least the TFIIIB subunit BDP1 [118]. Analysis of the 3D reconstruction of human τA showed that half of the cryo-EM density can be attributed to τA subunits TFIIIC35, TFIIIC63, and TFIIIC102 (equivalent to τ55, τ95, and τ131 in yeast), as shown in Figure 2.18 A. TFIIIC35 and TFIIIC63 form a heterodimer that binds to the C-terminal TPR domain of TFIIIC102, depicted in Figure 2.18 C. Thus, human τA resembles its yeast counterpart but lacks the phosphatase domain present in yeast τ55 (absent in hTFIIIC35) and the DNA binding domain of TFIIIC63.

Remarkably, the remaining half of the cryo-EM density is contributed by the C-terminal half of the τB subunit TFIIIC220, as seen in Figure 2.18. Insight into the interaction between the C-terminal portion of the human TFIIIC220 and the TFIIIC102 and TFIIIC63 subunits of τA was initially obtained when full human TFIIIC was treated with a poliovirus-encoded 3C protease, followed by specific antibody recognition targeting different fragments of TFIIIC220 [119]. The atomic model reveals a homeobox-like domain interacting with the TFIIIC35-TFIIIC63 heterodimer, and four WH domains (WH5 to WH8), with WH6 and WH8 making direct contact with TFIIIC102. TFIIIC220 occupies part of the TFIIIC102 binding interface that the TFIIIC63-DBD occupies in yeast, as shown in a previous study [37].
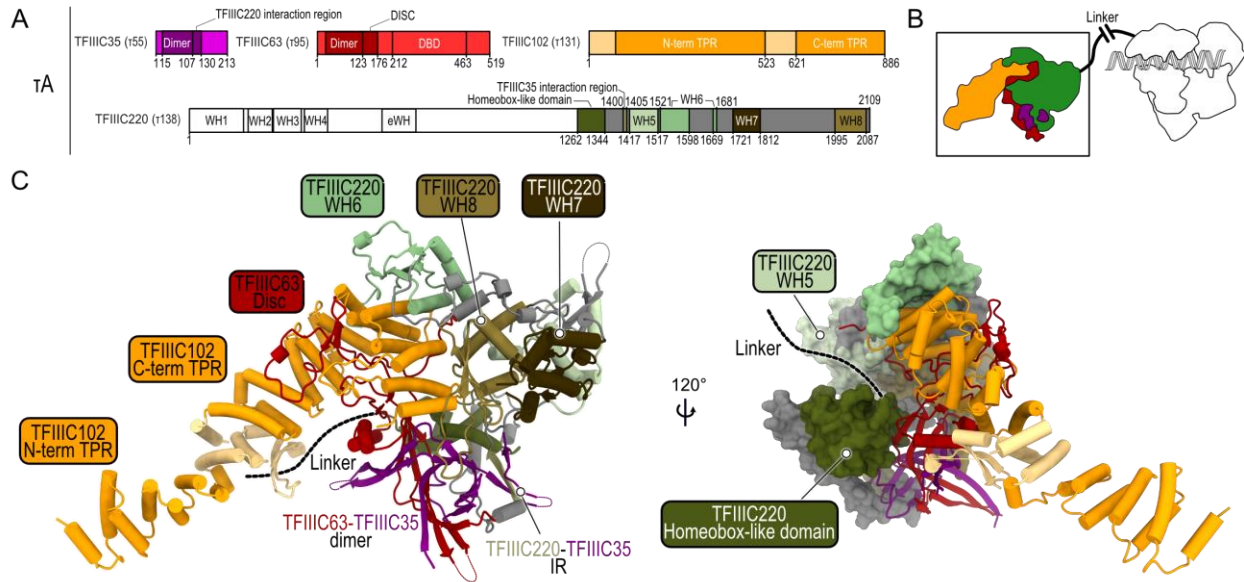
Figure 2.18 TFIIIC220 C-terminal region is part of the τA Subcomplex. A. Diagram illustrating the protein domains (shown in different colors) that comprise the stable τA subcomplex. For TFIIIC220, regions that were not modeled are depicted in gray. The interaction region is labeled as IR. **B.** Schematic depiction of τA (colored) and τB (white) connected by a flexible linker **C.** Atomic model of τA is presented, showing the C-terminus of TFIIIC220 in two forms: as a cartoon representation (left) and as a surface view (right). The flexible linker that connects τA and τB is indicated by black dashed lines. Panel A to C was adapted from reference [66].

The C-terminal half of TFIIIC220 is an integral component of human τA, suggesting a linkage between τA and τB via a roughly 550-residue linker between the N and C termini. To investigate this, we measured the distance between particles contributing to the τA and τB reconstructions using our cryo-EM data (refer to section 4.9.4). This measurement revealed a distinct distribution of particle distances, peaking at 200 Å, differing from a simulated set of randomly distributed particles (see figure 2.19 A). This suggests a stable yet flexible connection between τA particles and the tRNA gene-bound τB fraction, allowing TFIIIC to adapt to the varying distances between A-box and B-box in different human tRNA genes (as shown in figure 2.19 B).
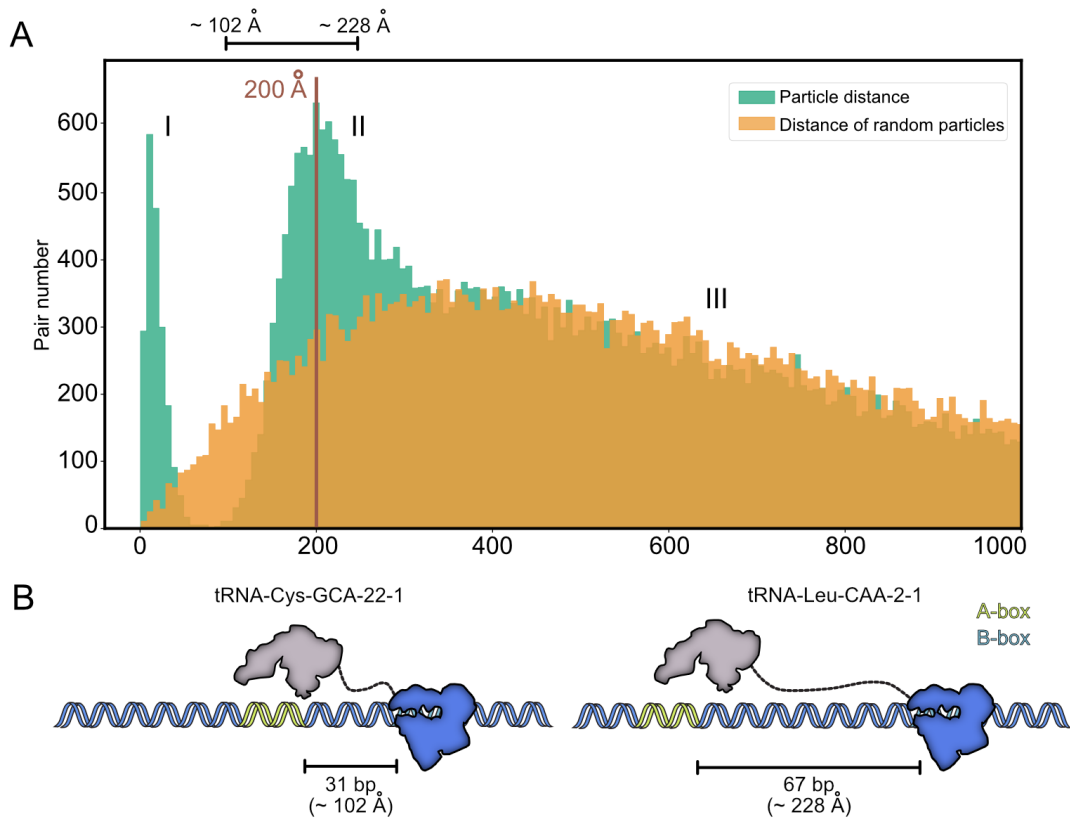
Figure 2.19. Connection of τA and τB via a flexible linker. A. Histogram comparing the distance distribution of τA and randomly generated particles. The histogram is accompanied by an annotation indicating the minimum and maximum distances observed between the A and B boxes in human tRNA genes. **B.** Schematics that illustrate the ability of TFIIIC to span both the shortest and longest distances between the A- and B-boxes found in human tRNAs. The shortest distance is exemplified by tRNA-Cys-GCA-22-1 and the longest by tRNA-Leu-CAA-2-1. In these schematics, τB is represented in blue, and τA is depicted in grey, highlighting the structural adaptability of TFIIIC in accommodating various tRNA gene configurations. Panel A to B was adapted from reference [66].

2.6.4 Model of TFIIIC promoter recognition

Based on our structural analysis, a multi-step mechanism for tRNA gene promoter recognition by TFIIIC, as illustrated in Figure 4, is suggested. The initial step involves the high-affinity sequence and shape recognition of the B-box promoter, facilitated by the assembly of different WH domains. WH2 and WH4 domains of TFIIIC220, part of the stable τB-core formed with TFIIIC90-TFIIIC110 and the TFIIIC220-WH1 domain, keep their conformation upon DNA binding. This suggests their role as a platform for initial DNA recognition. It is hypothesized that τB-WH2-WH4 assists in scanning the tRNA gene for the B-box promoter. Further stability in the TFIIIC-DNA interaction is provided by the WD40 domain in TFIIIC110, which binds to the DNA downstream of the B-box through a positively charged pocket. Following this, the WH3 and eWH domains engage the upstream half of the B-box, and collectively, the WH domains surround the upstream region along with the B-box motif. This comprehensive binding of the B-box motif by τB WH domains at both upstream and downstream sites anchors TFIIIC to the tRNA gene. The flexible 550-residue TFIIIC220 linker connecting τA and τB then facilitates the scanning of the tRNA gene for low-affinity A-box motifs, employing a fly-casting mechanism.
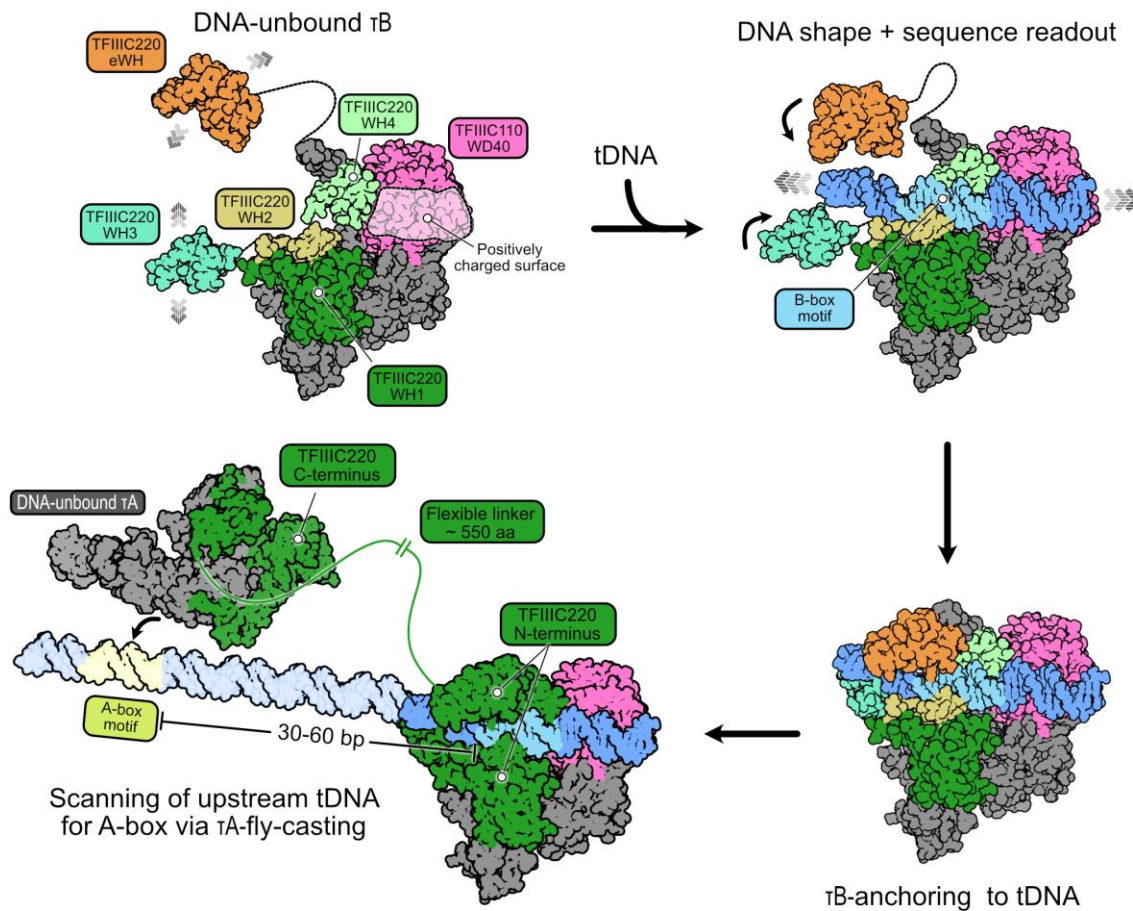
Figure 2.20. Proposed model for TFIIIC Promoter Recognition. The τB scaffold, comprised of TFIIIC110 and TFIIIC90 along with the WH1, WH2, and WH4 domains of TFIIIC220, forms a stable structure. The WH3 and eWH domains, in contrast, are more flexible. Initial DNA recognition is carried out by the WH2 and WH4 domains, followed by the WH3 and eWH domains, which engage the DNA through both shape and sequence recognition. Additionally, the positively charged pocket in the TFIIIC110-WD40 domain (shown with a transparent-white surface) contributes to stabilizing the DNA interaction. Once TFIIIC is securely anchored to the tRNA gene, the τA subcomplex, linked to τB by a flexible linker approximately 550 amino acids long, begins to search for the A-box motif. The A-box is located at a variable distance from the B-box, ranging between 30 and 60 base pairs. This search is conducted through a fly-casting mechanism, highlighting the dynamic and adaptable nature of TFIIIC in promoter recognition. Figure adapted from reference [66].

TFIIIC220, with a total of nine WH domains, plays several critical roles in TFIIIC-mediated promoter recognition. The N-terminal WH1 to WH4 and eWH domains are directly involved in B-box recognition (except WH1 which is indirectly involved), while the C-terminal WH5 to WH8 domains are essential components of τA, acting as a bridge between τA and τB. The flexible linker between τA and τB imparts TFIIIC with the necessary flexibility to adjust to the varying distances between A- and B-box motifs across different tRNA genes.

3. Cryo-EM analysis of yeast TFIIIC

Before the start of my PhD, Matthias Vorländer and Anna Jungblut, with Helga Grötsch's assistance, had already cloned and assembled all the yeast TFIIIC genes with their respective tags. A key point to mention is that for this project, I chose to work with the TFIIICΔ593 mutant, known for its higher affinity for the promoter A-box. This mutant had been previously published by our lab [37] and was available in our lab's database (refer to section 4.6.3 for more details). Although the purification strategy for these proteins was already standardized and documented, I opted to modify it. I simplified the original three-step purification process, which included His-tag, strep-tag, and anionic exchange, to just strep-tag and anionic exchange. This adjustment resulted in a similar yield and protein homogeneity as the published method, but significantly reduced the purification time from two days to just one day (see figure 3.1 for comparison).
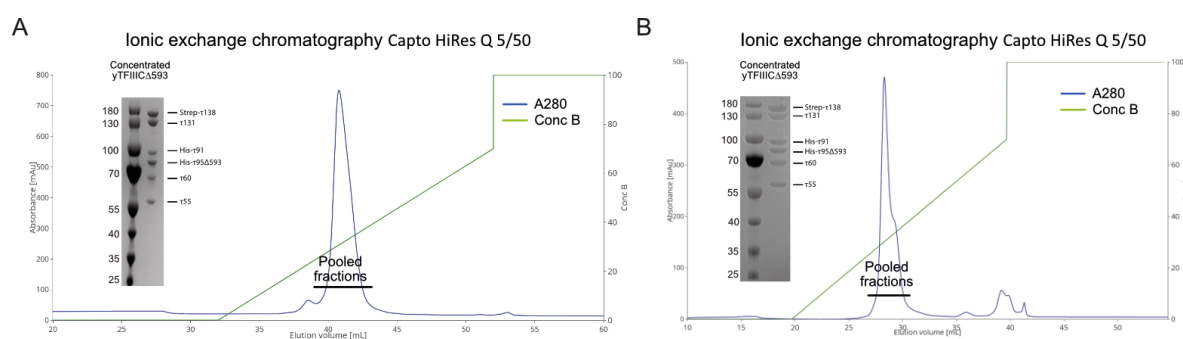
Figure 3.1 Comparison of the final step in yTFIIICΔ593 purification. A. Ion exchange chromatography used as the third step, with an SDS-PAGE gel stained with Coomassie Blue showing the concentrated yTFIIICΔ593 sample after pooling indicated fractions at the left. **B.** Ion exchange chromatography applied as the second step, with an SDS-PAGE gel stained with Coomassie Blue, presenting concentrated samples after pooling fractions containing yTFIIIC.

3.1 Optimization of buffers for Cryo-EM studies of the yTFIIICΔ593-DNA complex

For the cryo-EM studies of yTFIIICΔ593, the approach differed from that used with hTFIIIC. Instead of preparing fresh samples each time, I utilized previously purified and flash-frozen samples of yTFIIICΔ593 . Drawing on the knowledge acquired from my work on the hTFIIIC-DNA complex, I followed a similar method to ensure the stability of the yTFIIICΔ593 complex when combined with its tRNA gene. The His_tH(GUG)E2 gene, measuring 85 bp and having a molecular weight of 51 kDa, was selected for this purpose, in line with findings from our lab and outlined in a previous publication [37]. This gene was mixed with the yTFIIICΔ593 complex, and the resultant yTFIIICΔ593-DNA mixture was passed using a Zeba Spin desalting column (ThermoFisher Scientific) that was pre-equilibrated with a specific buffer. For detailed information on the buffer composition, please refer to section 4.10.1.

Following the buffer exchange, I performed mass photometry experiments to assess the influence of salt concentration on the stability of the yTFIIICΔ593 -DNA complex. This step was essential for identifying the ideal conditions for cryo-EM sample preparation. To see detailed buffer composition refer to section 4.4. To begin, I examined a sample without DNA. In this case, the molecular weight of the yTFIIICΔ593 complex was observed to be 547 ± 35 kDa, closely aligning with the theoretically calculated molecular weight of 520 kDa (refer to Figure 3.2, top left plot). Additionally, a distinct peak was observed,

representing a molecular weight slightly above half of the full complex (318-337 kDa). This could potentially indicate a dissociated TFIIIC, similar to results in a previous study employing scanning electron microscopy where TFIIIC particles varied in size from 510 to 670 kDa were directly observed, including smaller fractions approximately half that size [47]. Subsequently, I increased the salt concentration, starting from 150 mM and proceeding to 250 mM KCl in steps of 25 mM, and incorporated the tRNA gene into the experiments. This approach demonstrated that the yTFIIICΔ593 complex and DNA interaction remained stable up to 200 mM KCl. However, at higher concentrations (225 mM and 250 mM KCl), only the peak for yTFIIICΔ593 complex alone was evident. This suggests that at these higher salt concentrations, the yTFIIICΔ593 complex is unable to bind to its target tRNA gene (see figure 3.2 for more details). Interestingly, these results slightly deviate from previous studies on yTFIIIC, which indicated a complete absence of binding to its target RNA gene at and above 200 mM KCl in footprinting experiments [113].
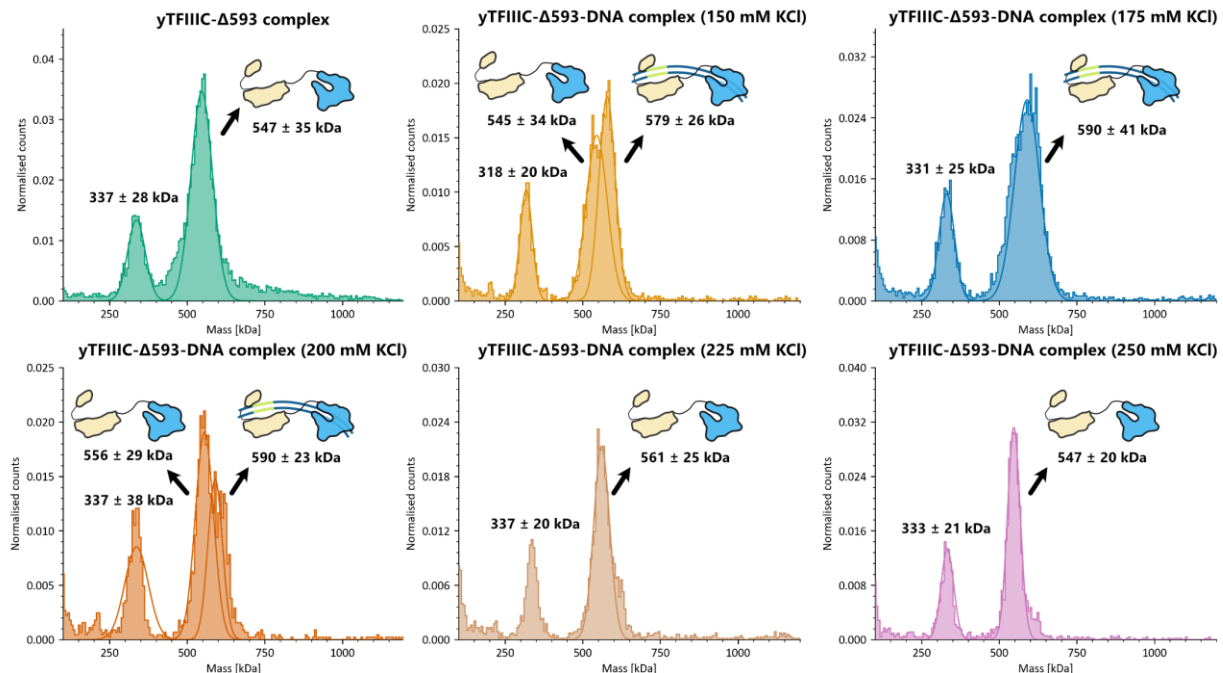
Figure 3.2 Mass photometry assessment of yTFIIICΔ593 -DNA complex stability across different ionic conditions. In the top left image, a single peak representing the TFIIICΔ593 complex in a DNA-free sample is seen. The top middle image shows overlapping peaks, indicating the presence of both TFIIICΔ593 and TFIIIC-Δ593-DNA complexes in a buffer containing 150 mM KCl. The top right image presents a distinct peak corresponding to the TFIIIC-Δ593-DNA complex in a buffer with 175 mM KCl. In the bottom left, overlapping peaks suggest a mixture of TFIIICΔ593 and TFIIIC-Δ593-DNA complexes in 200 mM KCl buffer. The bottom middle image displays a dominant peak for TFIIICΔ593 in 225 mM KCl, and the bottom right image also shows a single peak, indicating only the TFIIICΔ593 complex in a buffer with 250 mM KCl. Accompanying schematic diagrams next to each peak illustrate the TFIIICΔ593 and TFIIIC-Δ593-DNA complex, with molecular weights specified below each schematic. The consistent first peak across all conditions, corresponding to a 318-337 kDa protein, suggests potential disassembly of the full TFIIIC complex to one of the subcomplexes, either τA or τB. To find the full composition of the buffer used refer to section 4.4.

Mass photometry was critical in confirming the interaction between our complex and its target RNA gene across a range of salt concentrations. However, it did not identify the optimal salt concentration for this interaction. Therefore, I conducted parallel experiments, applying the same conditions used in mass photometry to cryo-EM sample preparation. The outcomes of these tests are summarized in Figure 3.3. Despite mass photometry suggesting 175 mM KCl as the best condition, due to a lack of two overlapping peaks unlike other conditions, incubating the yTFIIICΔ593 -DNA complex at this salt concentration led to aggregation, as shown in Figure 3.3 (left micrograph). Interestingly, the condition at the limit of salt concentration (200 mM KCl) for DNA binding detectable by mass photometry yielded a better outcome with more distinct individual particles and reduced aggregation, as seen in Figure 3.3 (middle micrograph). However, the 225 mM KCl condition, although indicating a single peak for the yTFIIICΔ593 complex in mass photometry experiments, displayed reduced complex formation and aggregation as observed in Figure 3.3 (right micrograph). These parallel experiments highlight the challenges in establishing suitable conditions for stabilizing the yTFIIICΔ593 -DNA complex for cryo-EM studies.
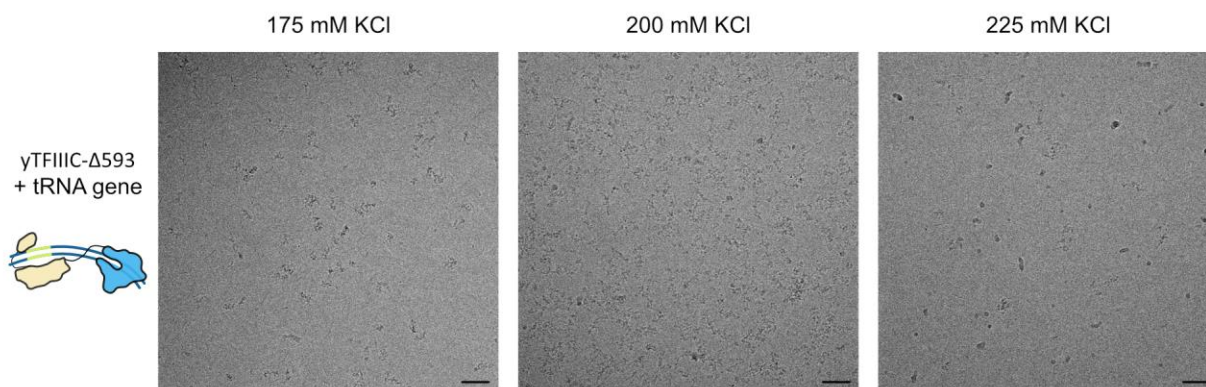
Figure 3.3. Optimization of buffer conditions for yTFIIICΔ593-DNA complex. The left micrograph, the complex is incubated with 175 mM KCl, shows significant aggregation. In the middle micrograph, the complex with 200 mM KCl displays reduced aggregation, and individual particles are visible. The right micrograph, with the complex in 225 mM KCl, reveals further aggregation, leading to an almost empty micrograph. Details on sample preparation are in section 4.10.1. Each micrograph is labeled with its corresponding salt concentration. The Ultrafoil R2/2 Au200 grid was used for all conditions, and a scale bar indicating 500 Å is included. Micrographs were taken at 92,000x magnification (pixel size of 1.566 Å).

3.2 Initial cryo-EM studies of the yTFIIICΔ593-DNA complex

Upon establishing that the yTFIIICΔ593-DNA complex remains stable in buffer containing 200 mM KCl after sample preparation for cryo-EM, I collected 19,047 movies using the Titan Krios microscope equipped with a K3 camera. The details regarding the components utilized are provided in Table 4.25, and the data acquisition process is thoroughly described in Table A3. The methodology used to produce the initial cryo-EM maps for both the τA-DNA and τB-DNA subcomplexes is outlined in Figure 3.4. During the initial data processing with WARP-picked particles, the yTFIIIC-DNA complex revealed three distinct sets of particles, in contrast to the two found in the hTFIIIC-DNA complex (refer to Figure 2.10). Unlike the hTFIIIC-DNA complex, which only showed τB-DNA subcomplex dimers, the yeast sample displayed both monomer and dimer forms (see Figure 3.4). Given the success of the data processing pipeline used for the hTFIIIC-DNA sample in achieving high-resolution structures, a similar approach was adopted for the yeast sample. This process involved using Topaz for training and particle picking, aimed at increasing the Number of particles specific to each distinct set. This was followed by iterative heterogeneous refinement and 2D classification, aimed at achieving better classification and isolating a homogeneous population of particles for each subcomplex. The final step was a non-uniform refinement using the best map obtained for each distinct class (see Figure 3.4).

After processing the yTFIIICΔ593-DNA dataset, it was observed that the τB-DNA monomer map reached a high resolution of 3.21 Å, setting it apart from the τB-DNA dimer, which displayed reduced resolution due to significant anisotropy, as seen in the direction distribution plot (Figure 3.4 – bottom left). This anisotropy may arise from the dimer particles' tendency to orient preferentially on the cryo-EM grid. For the τA-DNA subcomplex, the cryo-EM map showed a lower resolution, likely due to fewer particles being in a similar binding state, reflecting τA's dynamic interaction with DNA (Figure 3.5). In Figure 3.5, the 2D classes with different orientations indicate a region of the τA subcomplex maintaining steady contact with DNA, while others display variable binding and unbinding (red and yellow circles in figure 3.5 A). This observation aligns with the results from the 3D variability analysis performed in cryoSPARC, as shown in Figure 3.5 B, where τA is observed to intermittently engage and disengage with the DNA. Additionally, a density at the left end of the DNA, visible in the initial frame 0 volume, gradually disappears in subsequent frames. This transient density may correspond to the fuzzy density identified in the 2D classification (figure 3.5 A - white arrow).
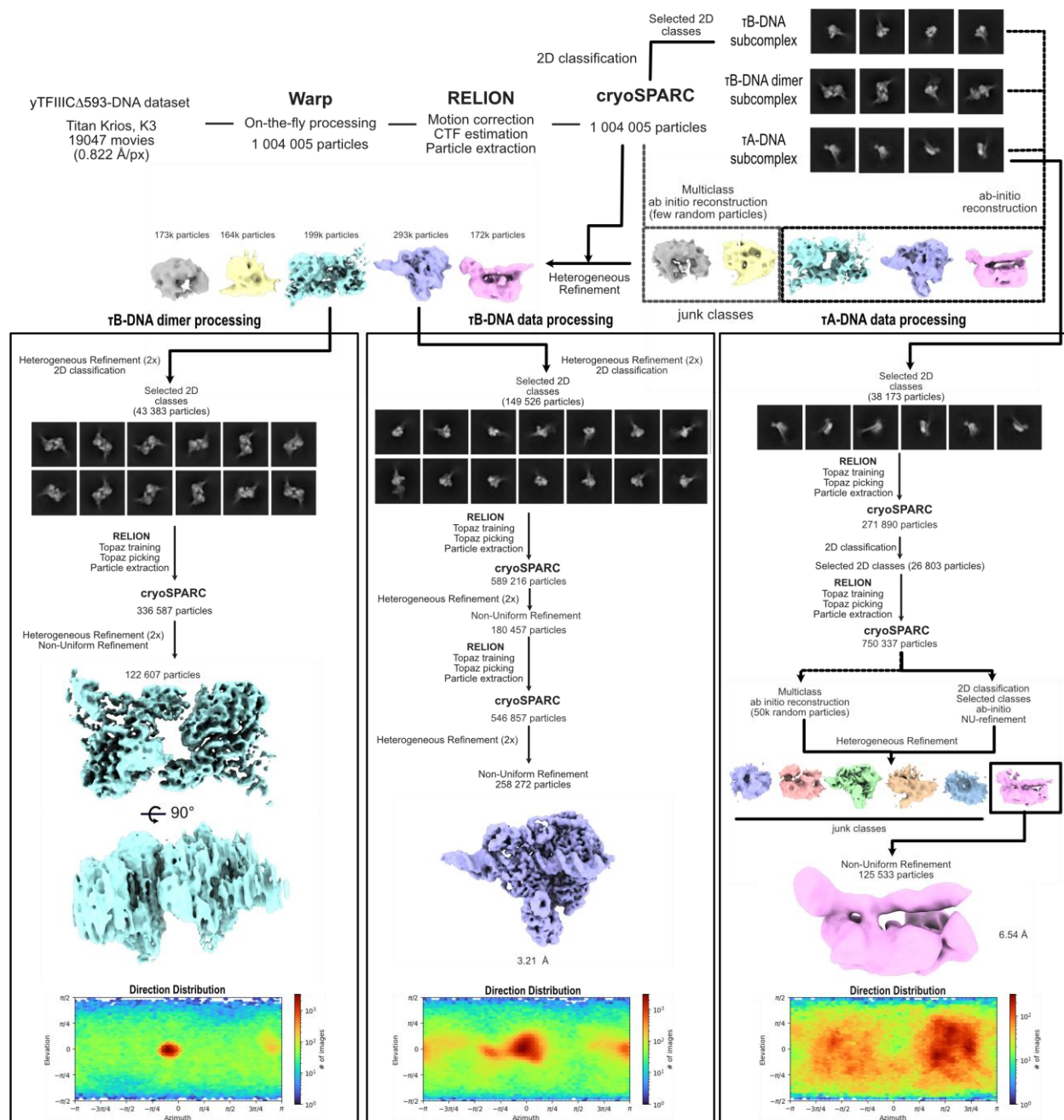
Figure 3.4: Workflow for the first cryo-EM dataset of the yTFIIICΔ593-DNA complex. The dataset, initially comprising 1,004,005 WARP-picked particles, was classified into τB-DNA dimer, τB monomer, and τA-DNA subcomplexes. Each class underwent multiple steps of particle training/picking, classification and refinement. The τB-DNA dimer, using the particles from the TOPAZ training and picking step, after two rounds of heterogeneous refinement, yielded a map containing 122,607 particles with preferred orientation (bottom left). The τB-DNA monomer class, following two rounds of TOPAZ training and picking, resulted in a map containing 258,272 particles with a resolution of 3.21 Å (bottom middle). Similarly, the τA-DNA subcomplex, following two rounds of TOPAZ training and picking and subsequent heterogeneous refinement resulted in a 125,533 particle map with a resolution of 6.54 Å (illustrated at bottom right). The particle distribution for each distinct class is displayed at the lower section of their respective data processing pipeline.

43
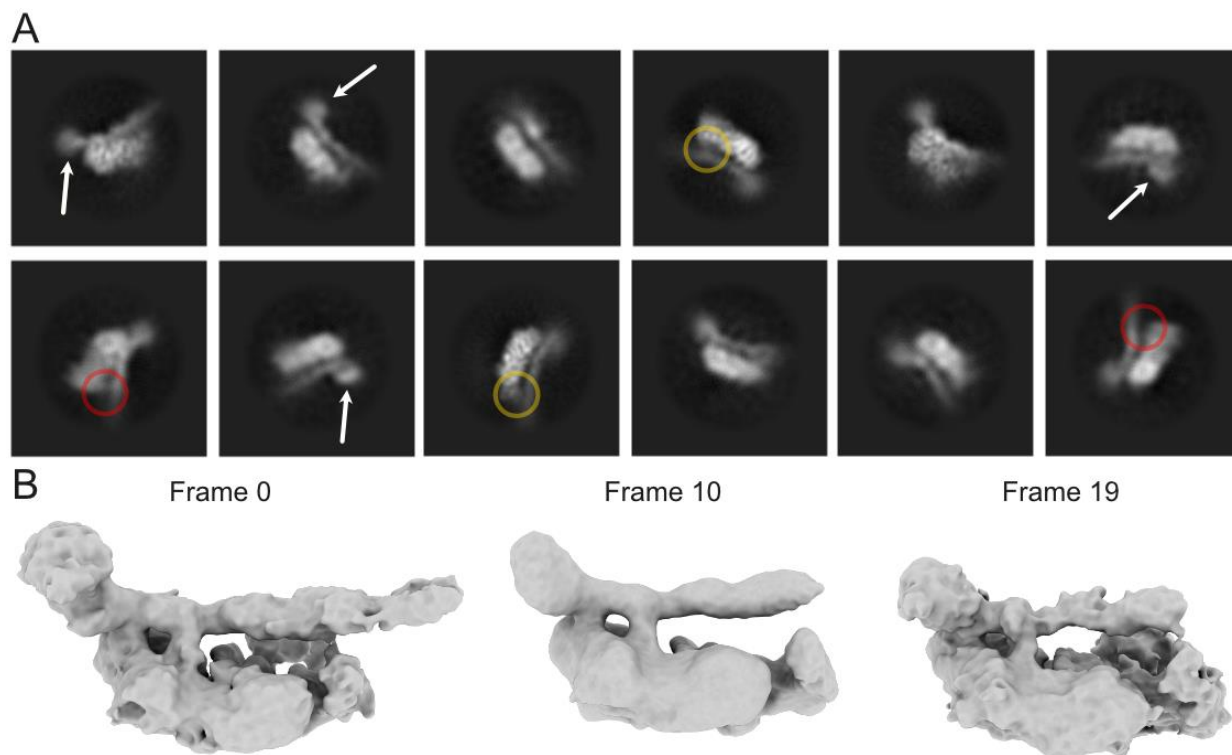
Figure 3.5: Dynamic interaction between τA and DNA. **A.** Representative 2D class averages from the first cryo-EM dataset show regions where τA engages with DNA (highlighted by yellow circles) and disengages (indicated by red circles). Also a dynamic fuzzy region at one of the DNA termini is also visible (denoted by white arrows). **B.** 3D variability analysis volumes (component 0) obtained using cryoSPARC applying a resolution filter of 8 Å reveal an extra density at the left end of the DNA (frame 0) that is not present in the other frames. Additionally, partial and periodic association and dissociation of τA with the DNA is shown (frame 0 – association, frame 10 – dissociation, frame 19 – re-association).

The investigation into the oligomerization of hTFIIIC and its DNA complex was constrained by the limited sample quantity, preventing a thorough exploration of this aspect. To address this gap, the yTFIIICΔ593 complex was used. Notably, when examining the first cryo-EM dataset, it is important to recognize that, despite the use of a lower protein concentration during sample preparation for the human complex (1.53 µM) compared to the yeast complex (1.97 µM), monomer subcomplexes were consistently not detected in the human sample after data processing. This absence, even at lower concentrations, suggests that the differences in concentration between the yeast and human samples are not the only factor influencing dimer formation within these complexes. Furthermore, initial mass photometry experiments for both human (refer to Figure 2.7) and yeast (Figure 3.2) TFIIIC did not show any dimer formation. This observation led to the hypothesis that oligomerization might be concentration-dependent, which explains why it was not detected at the lower concentrations used in these experiments (20 nM), in contrast to the higher concentrations (1.53 µM for human and 1.97 µM for yeast) employed during sample preparation.
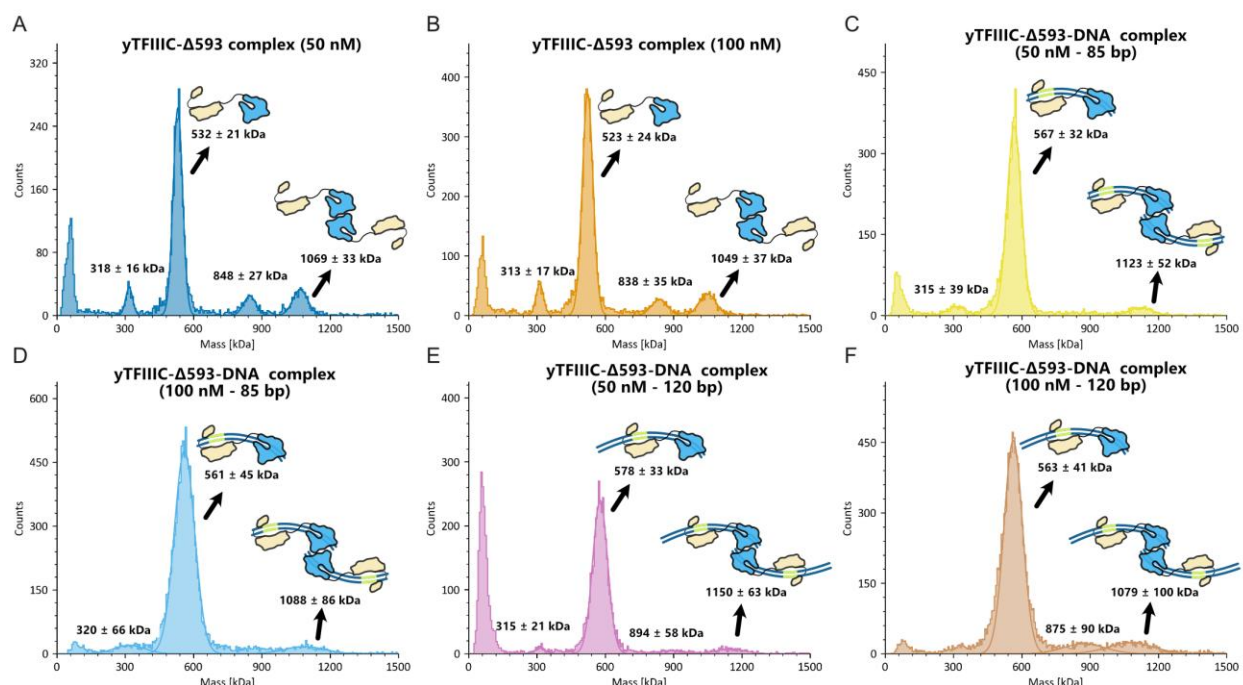
Figure 3.6: Mass Photometry analysis of the yTFIIICΔ593 and yTFIIICΔ593-DNA complex dimerization. A. & B: Mass distribution histograms of yTFIIICΔ593 complex without DNA at 50 nM (A) and 100 nM (B) concentrations. At 50 nM, the main peak is at 532 ± 21 kDa (monomer), with higher peaks at 848 ± 27 kDa and 1069 ± 33 kDa (dimer). At 100 nM, the main peak is at 523 ± 24 kDa (monomer), with higher peaks at 838 ± 35 kDa and 1049 ± 37 kDa (dimer).**C & D:** Mass distribution histograms of yTFIIICΔ593 complex with DNA (85 bp) at 50 nM (C) and 100 nM (D) concentrations. At 50 nM, the main peak shows 567 ± 32 kDa (monomer-DNA) and a higher peak at 1123 ± 52 kDa (dimer-DNA). At 100 nM, the main peak is 561 ± 45 kDa (monomer-DNA), with a higher peak at 1088 ± 86 kDa (dimer-DNA). **E & F:** Mass distribution histograms of yTFIIICΔ593 complex with DNA (120 bp) at 50 nM (E) and 100 nM (F) concentrations. At 50 nM, the main peak is 578 ± 33 kDa (monomer-DNA), with higher peaks at 894 ± 58 kDa and 1150 ± 63 kDa (dimer-DNA). At 100 nM, the main peak is 563 ± 41 kDa (monomer-DNA), with higher peaks at 875 ± 90 kDa and 1079 ± 100 kDa (dimer-DNA). Accompanying schematic diagrams next to each analyzed peak illustrate the TFIIICΔ593 and TFIIIC-Δ593-DNA complex in monomer and dimer forms, with molecular weights specified below each diagram. The consistent first peak across all conditions, corresponding to a 313-320 kDa protein, suggests potential disassembly of the full TFIIIC complex into one of its subcomplexes, either τA or τB. Refer to Section 4.4 for the full composition of the buffer used.

Standard practices for MP experiments generally suggest using protein or complex concentrations within the 10 nM to 50 nM range[120]. Recognizing that initial MP assays of both hTFIIIC and yTFIIICΔ593, with and without DNA, were conducted at 20 nM (refer to Section 4.4 for more details), exploring higher concentrations was considered valuable to potentially reveal additional oligomerization states in the absence or presence of DNA (see Figure 3.6). When the concentration of yTFIIICΔ593 complex alone was increased to 50 nM and 100 nM (Figure 3.6 A and B, respectively), two notable new peaks emerged: one around 840 kDa and another approximately at 1060 kDa. It is important to note that there were no significant differences in the counts of these two higher peaks between the 50 nM and 100 nM conditions. Intriguingly, the first peak corresponds to the combined molecular weight of the yTFIIICΔ593 complex and the smaller unidentified peak from Figure 3.2, presumably either τA or τB, as also seen in these MP experiments. While the exact identity of this smaller peak is yet to be confirmed,

there is a reasonable possibility it represents the τB subcomplex. This hypothesis aligns with the initial cryo-EM data, which suggests that the interaction between two yTFIIICΔ593 complexes is mediated by the τB complex in higher concentration conditions. If the unidentified smaller peak is indeed τB, it is likely to interact with the yTFIIICΔ593 complex, resulting in the appearance of the 840 kDa peak. Further experiments, however, are necessary to conclusively identify this subcomplex. The highest peak corresponds to the yTFIIICΔ593 dimer. The same results were observed when yTFIIICΔ593 complex was incubated with the His_tH(GUG)E2 gene fragment of 85 bp (Figure 3.6 C and D) and 120 bp (Figures 3.6 E and F) lengths. The presence of dimer in the sample with and without DNA indicates that dimerization is not DNA-dependent, confirming the findings obtained with the hTFIIIC sample (for differences between the DNA oligo nucleotides used, refer to Table 4.25).

3.3 High resolution cryo-EM structures of the yTFIIICΔ593-DNA complex

When I began processing the first yTFIIICΔ593-DNA dataset, I also explored stabilizing this complex with additional transcription factors, such as fpt1 and components of TFIIIB, employing the same approach used for the yTFIIICΔ593-DNA complex. This exploration included confirming the interaction of the yTFIIICΔ593-DNA complex with fpt1 and with either the complete yTFIIIB or the yBTB (Brf1-TBP fusion protein), as evidenced by EMSA or mass photometry, detailed in figure 3.18. Following these confirmations, several datasets were collected. For details on the specific components used in each dataset, please refer to table 4.25. Additionally, information regarding the data acquisition parameters on the Titan Krios microscope can be found in section 4.10.2.

The steps involved in processing all datasets to improve the resolution for the τB-DNA monomer complex, compared to the resolution of 3.21 Å obtained from the initial dataset, are detailed in figure 3.7. The preliminary analysis of the new datasets revealed the presence of all known subcomplexes: τA-DNA, τB-DNA monomer, and τB-DNA dimer. This finding was initially viewed as a positive indication that the additional transcription factors might be contributing to the stabilization of the yTFIIICΔ593-DNA complex during sample preparation. This hypothesis was based on the observed instability of the yTFIIICΔ593-DNA complex when prepared alone, particularly at salt concentrations below 200 mM KCl. This instability was clearly showed in Figure 3.3. Notably, the samples in this study were prepared using a lower range of salt concentrations, as detailed in Table 4.25, further suggesting the stabilizing role of the additional factors. Upon closer inspection of the final maps from each dataset, no extra densities were observed that would suggest the presence of fpt1 or the components of TFIIIB. This lack of extra density could imply either disassembly of the transcription factors from the yTFIIICΔ593-DNA complex during cryo-EM sample preparation or, if the transcription factors are bound to the τA-DNA or τB-DNA subcomplexes, their potential flexibility might prevent their resolution in the final maps

In summary, all datasets were processed using consistent procedures. The initial particles picked by WARP were subjected to heterogeneous refinement, using a volume of the τB-DNA complex, obtained by ab-initio reconstruction, and several decoy volumes to filter out junk particles. The Number of heterogeneous refinement rounds varied among the datasets, ranging from one (as in dataset 2) to up to four (as in dataset 4). A non-uniform refinement step was applied to the best map from each dataset to evaluate its resolution. The particles from these refined maps were then merged and went through two more rounds of heterogeneous refinement. This process yielded a final map with a resolution of

2.63 Å, composed of 833,266 particles selected from the initial pool of 1,280,178 merged particles. The quality of this map was verified by a Fourier Shell Correlation (FSC) curve, which displays the achieved resolution (Figure 3.7 – bottom left). Additionally, a particle distribution chart is included in figure 3.7 (bottom right), showing the distribution of the particles that contributed to this map. It should be noted that attempts to obtain a high-resolution map for the τB-DNA dimer complex were not pursued due to the preferred orientation of these particles across all collected datasets. This issue persisted despite experimenting with various salt conditions or the presence of different transcription factors added to the yTFIIICΔ593-DNA complex, which were strategies that could have influence the τB-DNA dimer's interaction with the air-water interface.

In the initial dataset for the yTFIIICΔ593-DNA complex, the τA-DNA subcomplex was resolved to a lower resolution of 6.54 Å, as shown in Figure 3.4 (Bottom right). To improve the resolution of the τA-DNA complex, additional datasets of the yTFIIICΔ593-DNA complex with various transcription factors were utilized to increase the pool of τA-DNA particles. The goal of using a larger set of initial particles was to aim, after an extensive classification procedure, at finding a homogeneous population that represents a specific binding state of the τA-DNA subcomplex. This τA-DNA interaction has been demonstrated to be very dynamic, as ilustrated in Figure 3.5 A and B. This dynamic interaction is likely one of the reasons why achieving a high-resolution τA-DNA map with only a single dataset proved challenging. The data processing began with on-the-fly particle picking using WARP, followed by motion correction and CTF estimation in RELION (see figure 3.8). The particles were then imported to cryoSPARC for further initial analysis. To increase the initial particle count from WARP, two additional rounds of particle training and picking with TOPAZ were performed for each dataset, except for dataset 4, which was subjected to just one round. After the τA-DNA particles were identified, cryoSPARC was used to generate an ab-initio map. This map, alongside decoy volumes, was used to carry out heterogeneous refinement, with the aim of removing particles not associated with the τA-DNA subcomplex. The volumes obtained from this heterogeneous refinement step in dataset1 were applied across datasets 2 to 4 for an identical heterogeneous refinement process. For dataset 5, new volumes were specifically generated using an ab-initio reconstruction approach. Following this, particles from each dataset after the last heterogeneous refinement were further cleaned through 2D classification, discarding only obvious junk classes (i.e. ethane contamination). These particles from all the datasets were merged and processed through four additional rounds of heterogeneous refinement, leading to a map consisting of 114,321 particles. Following non-uniform refinement, this map achieved a resolution of 3.73 Å.

Dataset1
(yTFIIICΔ593-DNA)

Dataset2
(yTFIIICΔ593-TFIIIB-fpt1-DNA)

Dataset3
(yTFIIICΔ593-fpt1-DNA)

Dataset4
(yTFIIICΔ593-TFIIIB-fpt1-DNA)

Dataset5
(yTFIIICΔ593-Brf1-TBP-DNA)

19047 movies
(0.822 Å/px)

16289 movies
(0.822 Å/px)

11644 movies
(0.822 Å/px)

16169 movies
(0.822 Å/px)

15614movies
(0.822 Å/px)

Warp
On-the-fly processing
1 004 005 particles

Warp
On-the-fly processing
960 668 particles

Warp
On-the-fly processing
1 344 141 particles

Warp
On-the-fly processing
1 566 773 particles

Warp
On-the-fly processing
1 129 934 particles

RELION
Motion correction
CTF estimation
Particle extraction

RELION
Motion correction
CTF estimation
Particle extraction

RELION
Motion correction
CTF estimation
Particle extraction

RELION
Motion correction
CTF estimation
Particle extraction

RELION
Motion correction
CTF estimation
Particle extraction

cryoSPARC

cryoSPARC

cryoSPARC

cryoSPARC

cryoSPARC

2D classification

Selected 2D classes

τB ab-initio reconstruction

Multiclass ab initio reconstruction

Multiclass ab initio reconstruction

Multiclass ab initio reconstruction

Multiclass ab initio reconstruction

Multiclass ab initio reconstruction

Heterogeneous Refinement

Heterogeneous Refinement

Heterogeneous Refinement

Heterogeneous Refinement

Heterogeneous Refinement

199k particles — 281k particles — 185k particles — 252k particles

198k particles — 190k particles — 118k particles — 254k particles

173k particles — 164k particles — 199k particles

164k particles — 160k particles — 326k particles — 186k particles

244k particles — 176k particles — 128k particles — 168k particles

293k particles — 172k particles

156k particles

207k particles — 187k particles

183k particles — 198k particles

159k particles — 125k particles

376k particles

176k particles

Heterogeneous Refinement (2x)
Non-Uniform Refinement

Non-Uniform Refinement

Non-Uniform Refinement

Heterogeneous Refinement (4x)
Non-Uniform Refinement

Non-Uniform Refinement

3.28 Å
191 618 particles

3.03 Å
281 077 particles

3.08 Å
326 061 particles

3.00 Å
226 854 particles

3.38 Å
254 568 particles

Heterogeneous Refinement (2x)
Non-Uniform Refinement
DeepEMhancer

cryoSPARC
FSC validation

cryoSPARC
Direction distribution

2.67 Å
833 266 particles

Fourier Shell Correlation — 1/resolution — 0.143 — 2.67 Å

Elevation — Azimuth — # of images

Figure 3.7. General workflow processing all the cryo-EM datasets to obtain a high resolution cryo-EM map of τB-DNA subcomplex. After the initial on-the-fly data processing in WARP, which included particle picking for each dataset, the particle coordinates were imported into RELION. Simultaneously, data preprocessing steps were carried out in RELION. Following this, the extracted particles were imported to cryoSPARC. In cryoSPARC, the τB-DNA subcomplex particles were identified, and an ab-initio reconstruction step was performed. Additionally, a multiclass ab initio reconstruction was applied to generate decoy volumes to sort out junk particles. Several rounds of heterogeneous refinement were performed to find a τB-DNA homogenous population of particles. Prior to combining the final particles from all five datasets, a non-uniform refinement step was implemented. These particles (after merging) were again further cleaned by 2 rounds of heterogeneous refinement to reach a final set of 833 266, which gave a map of 2.67 Å resolution. The FSC curve of this map (bottom left) and the direction distribution of these particles (bottom right) are included. Representative 2D class averages, 3D classifications from heterogeneous refinement, and non-uniform refinement maps of τB-DNA along with junk volumes are also displayed.
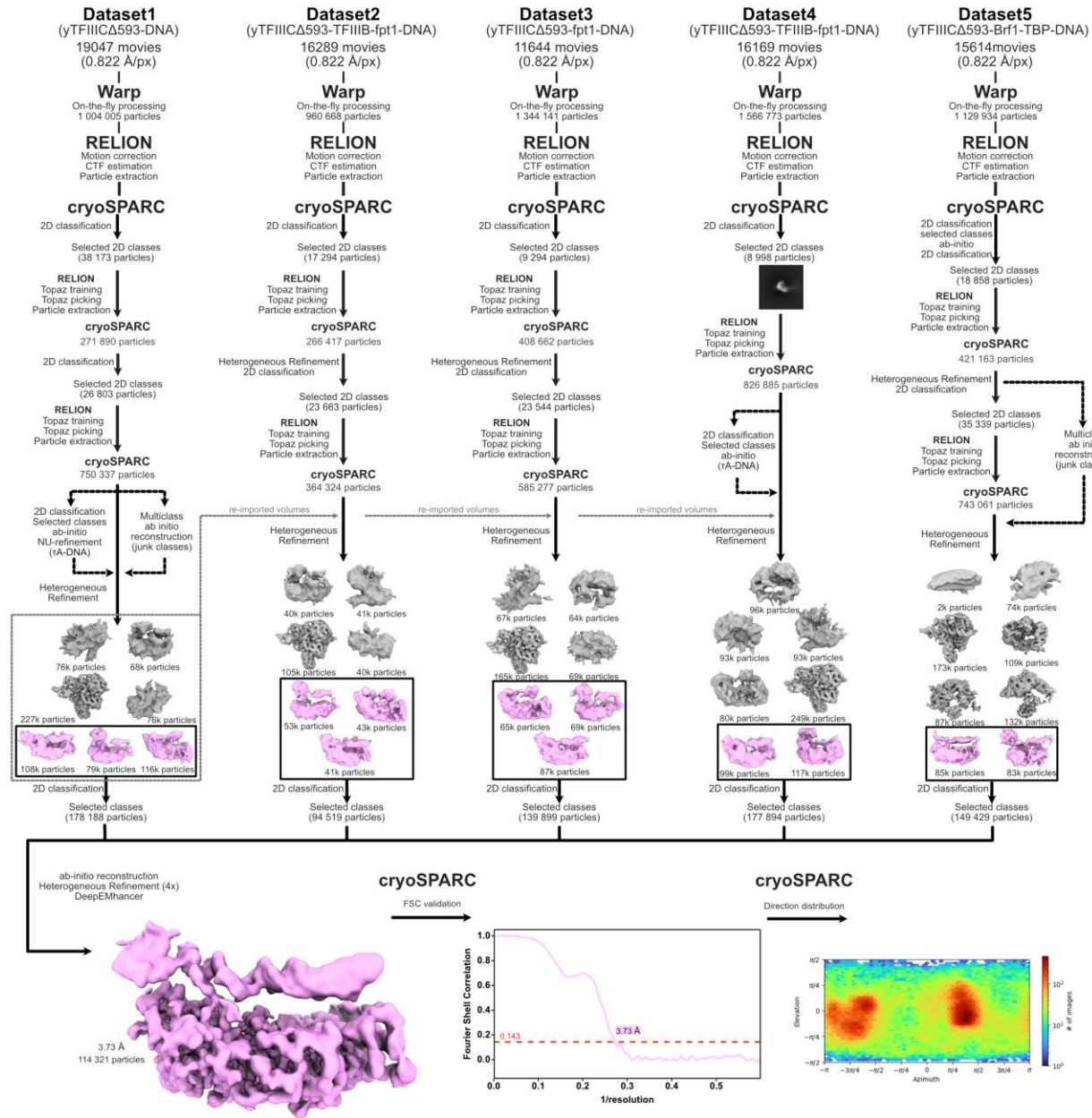
Figure 3.8. General workflow applied to all the cryo-EM datasets to produce a high resolution cryo-EM map of τA-DNA subcomplex. The process started with on-the-fly processing and particle picking in WARP. Next, preprocessing steps such as motion correction and CTF estimation were carried out in RELION. The extracted particles were then imported into cryoSPARC for further processing. To increase the initial particle Number from WARP, two rounds of TOPAZ training and picking were performed for each dataset, except for dataset 4, which underwent only one round. Following the identification of τA-DNA particles, an ab-initio reconstruction was conducted. This was followed by creating multiple volumes (junk classes) for subsequent heterogeneous refinement. After the final heterogeneous refinement job in each dataset, the particles are further cleaned using 2D classification before merging, where only obvious junk classes were discarded. The merged particles then undergo four rounds of heterogeneous refinement. This last step yielded a map containing 114,321 particles, which after non-uniform refinement reached 3.73 Å resolution. The figure also includes an FSC curve of the final volume and a plot showing the direction distribution of the particles that contributed to this map.

3.3.1 Structures of yeast τA-DNA and τB-DNA subcomplexes

Upon obtaining the final cryo-EM maps for the τA-DNA and τB-DNA subcomplexes, model building was conducted, as shown in figure 3.10 for model validation. The domain architecture of the yeast TFIIIC subunits is depicted in Figure 3.9 A. The subunits of yeast TFIIIC are divided into two groups: τ138 C-terminus, τ131, τ95Δ593, and τ55 form the τA subcomplex, while τ138 N-terminus, τ91, and τ60 constitute the τB subcomplex. The cloning and subsequent characterization of these individual subunits have clarified the composition of the τA and τB subcomplexes, as detailed in references ([50]–[55], [79], [121]). However, aligning with human TFIIIC studies (refer to section 2.6.1), it is proposed that the functionally active τA subcomplex should include τ138's C-terminus, with only its N-terminus being part of the τB subcomplex. This refined understanding offers a clearer perspective on the functional organization within the yeast TFIIIC complex.
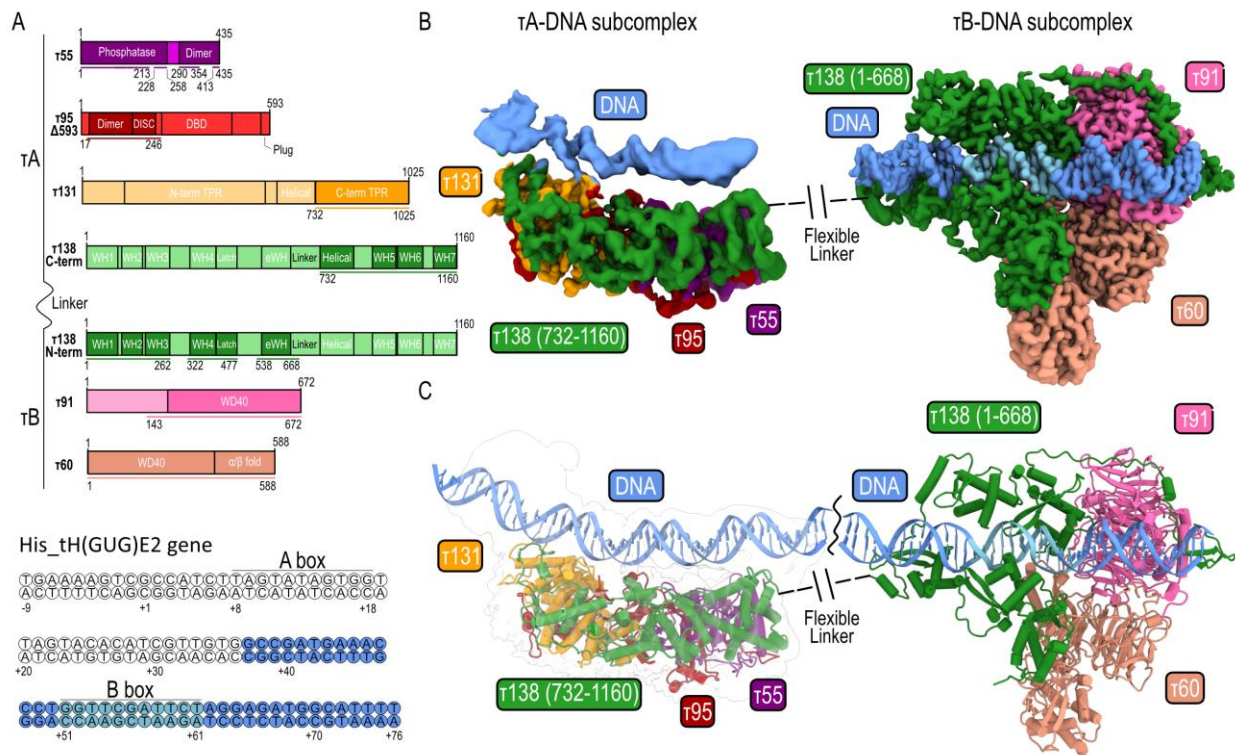
Figure 3.9. Cryo-EM maps and models of τA-DNA and τB-DNA subcomplexes. A. At the top, domain architecture of yTFIIIC subunits, categorized into either τA or τB subcomplexes, with a colored line indicating the extent of each subunit that could be structurally resolved. Below this, the sequence of the 85bp His_tH(GUG)E2 gene is displayed. The colored nucleotides indicate successful annotation without ambiguity due to high resolution of the map (τB-DNA subcomplex). **B.** Cryo-EM maps of the yTFIIICΔ593-DNA complex. The maps were postprocessed by DeepEMhancer [122] **C.** The cryo-EM models of the yTFIIICΔ593-DNA complex. The τA-DNA model was placed within its local resolution map, applying a lower threshold compared to the deepEMhancer map used in B. This was performed to highlight the density where the additional DNA structure was built.

Despite the τB-DNA subcomplex having a resolution of 2.67 Å, which was sufficient to assign the downstream region of the His_tH(GUG)E2 DNA oligonucleotide (40 base pairs), the τA-DNA subcomplex, resolved at 3.73 Å, had certain limitations. This lower resolution in the τA-DNA subcomplex prevented unambiguous nucleotide annotation. Consequently, this limitation also affected the accurate localization of the A-box within the density. However, utilizing the local resolution map at this lower threshold enabled the construction of the remaining segment of the His_tH(GUG)E2 gene (45 base pairs), as seen in figure 3.9 C. This limitation explains why, in Figure 3.9 A (bottom), only the confidently assigned downstream region of the His_tH(GUG)E2 gene is colored, while the remainder of the sequence is drawn in white.
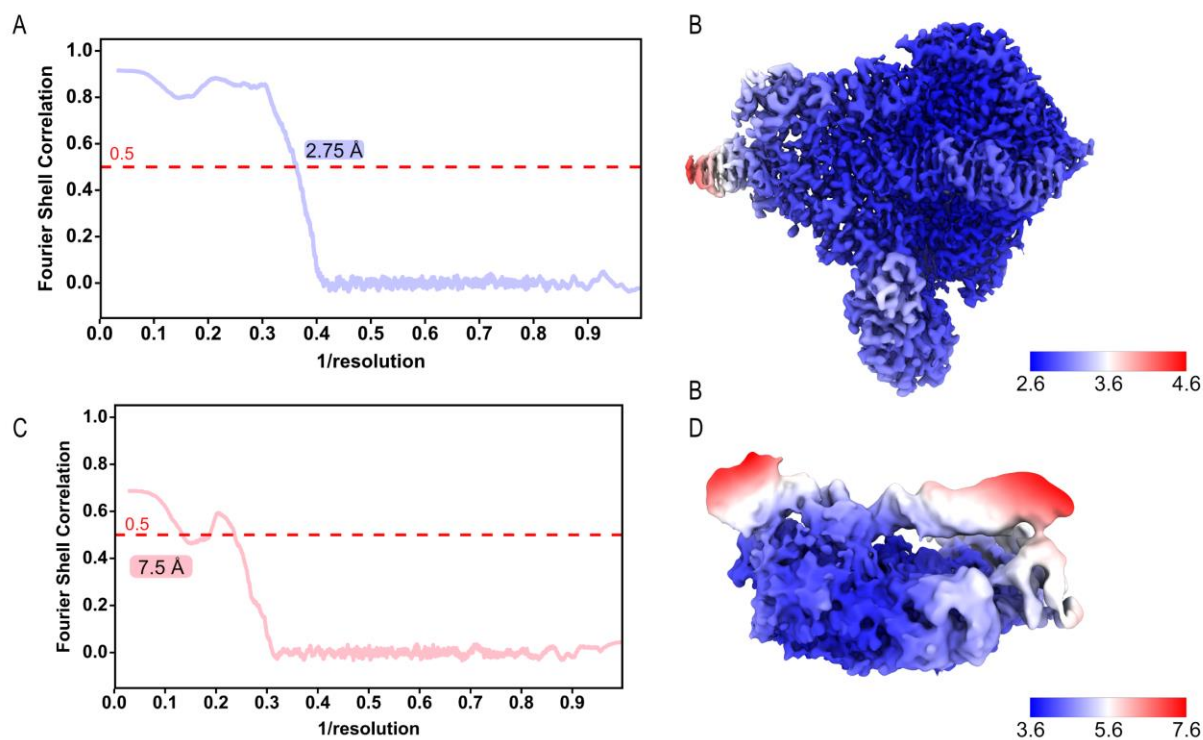
Figure 3.10. Model validation of the yTFIIICΔ593-DNA complex A. a Fourier Shell Correlation (FSC) is computed between the τB-DNA map and its corresponding model to assess the fit of the model to the map, using an FSC criterion of 0.5. **B.** The local resolution for the τB-DNA subcomplex map, determined using RELION software. **C.** Similarly an FSC curve between the τA-DNA map and its model, evaluating the map-to-model fit with an FSC criterion of 0.5. **D.** The local resolution for the τA-DNA subcomplex map, calculated using RELION.

The dimerization of yeast TFIIIC has been demonstrated through biophysical experiments like mass photometry, as well as in early cryo-EM studies (see section 3.2). In human TFIIIC, the interaction responsible for oligomerization is mediated by the TFIIIC110 subunit, corresponding to τ91 in yeast (refer to figures 2.11 – 2.13). However, in yeast, this interaction appears more dynamic, as shown in Figure 3.11. The initial dataset, along with subsequent datasets incorporating various transcription factors and buffer conditions, revealed a preferred orientation of τB-DNA dimer particles towards the air-water interface. This orientation preference prevented obtaining a high-resolution cryo-EM map of the oligomeric state. Notably, this issue was not observed with the τB-DNA subcomplex as a monomer.

Further analysis of the 2D classification elucidated two distinct forms of interaction between yeast τB-DNA monomers, as detailed in Figure 3.11 A. The first interaction type, termed the "thumb-knuckle" dimer, is characterized by the thumb of one hand in a thumbs-up position touching the knuckle of another hand in a thumbs-down position, which is visualized in Figure 3.11 A and specifically identified by a representative 2D class encircled in red (top left). The second interaction, termed the "knuckle-knuckle" form, resembles two hands, one in a thumbs-up and the other in a thumbs-down position, touching at the knuckles, as shown in Figure 3.11 A (top middle). However, due to the preferred orientation of the particles, constructing a reliable 3D map of these dimers was not feasible. Among the dimers observed, the "thumbs-knuckle" type was most prevalent across all datasets. Rigid body fitting of two τB-DNA monomer models into this map indicated that the τ60 subunit (TFIIIC90 in humans) is crucial for the "thumb-knuckle" dimer formation. Conversely, the "knuckle-knuckle" dimer is thought to be primarily mediated by the τ91 subunit as in human τB-DNA dimer (see Figure 3.11 B). This variation in dimerization forms highlights the adaptability of yeast TFIIIC dimerization, contrasting with its human counterpart.
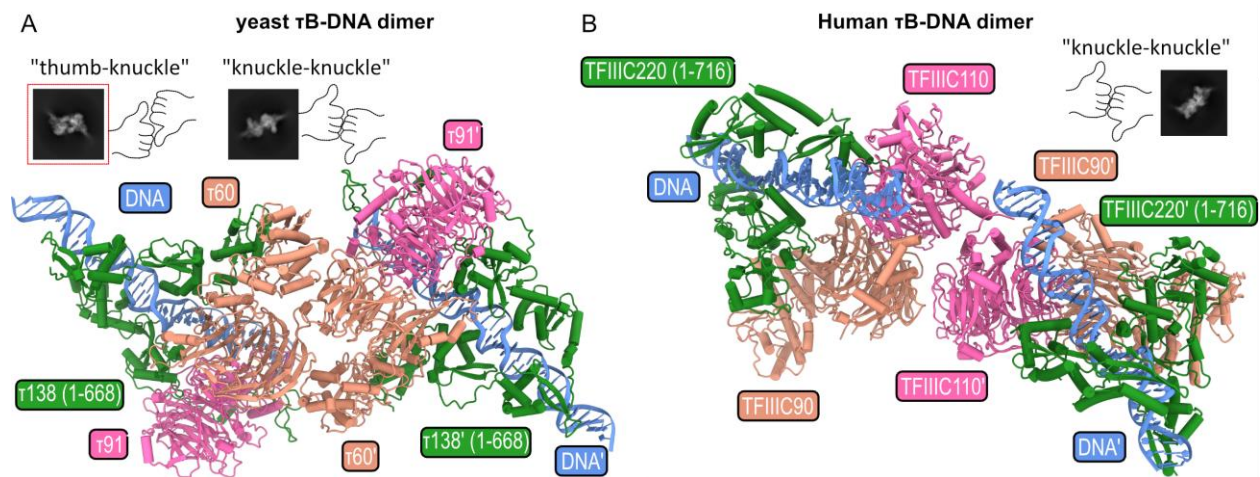
Figure 3.11. Dimerization differences between yeast and human τB-DNA subcomplexes. A. In yeast, dimerization of the τB-DNA complex might occur in two forms. One form, referred to as the "thumb-knuckle," involves engagement between each monomer through the τ60 subunits (TFIIIC90 in humans). The other form, termed "knuckle-knuckle," is mediated through the τ91 subunits (equivalent to TFIIIC110 in humans). **B.** Human τB-DNA dimerization predominantly takes place through the TFIIIC110 subunit (τ91 in yeast), following the "knuckle-knuckle" form. The yeast τB-DNA dimer model, representing the "thumb-knuckle" type, was obtained by oerforming rigid body fitting of two τB-DNA monomers into the corresponding cryo-EM map. Representative 2D classes for each dimer type are shown, with the "thumb-knuckle" form in yeast being highlighted in a red dashed square, indicating the dimerization type represented in the 3D model. For the human τB-DNA dimer model, coordinates were retrieved from the Protein Data Bank (PDB ID: 8CLJ).

3.3.2 Interaction of yeast τB subcomplex with tRNA gene

The yeast τB core, consisting of τ60 and τ91 subunits, closely resembles the human TFIIIC core composed of TFIIIC110 and TFIIIC90, detailed in section 2.6.2. A similar structure can be found of a truncated version of this core, involving τ60 and Δτ91 (aa 159-672), using X-ray crystallography [59]. Notably, this truncated form exhibited no capacity for DNA binding, showing the critical role of τ91 N-terminal region (aa 1-158) in DNA interaction. In the cryo-EM structure, it is observed that the subunit τ91 binds to the downstream region of the B-box, a process that might involve sequence specificity, as evidenced by interactions of K591 and N628 with the specific bases (see figure 3.12 – bottom right). Despite these amino acids being present in the τ60/Δτ91 structure, the absence of segments from aa 142 to 158, particularly K153 and R155, which interact with phosphate backbone of the DNA in the τB-DNA structure, might be critical for stabilizing this interaction. In the human TFIIIC, the TFIIIC110 subunit with its positively charged pocket is hypothesized to stabilize the interaction τB-DNA (figure 2.16). An study, where the subunit τ91 was cloned and characterized for the first time, found that this subunit cooperates with τ138 for DNA binding [23]. Furthermore, τ60 in yeast also contributes to DNA binding, engaging with the phosphate backbone of the B-Box motif through R11 and S362, as seen in figure 3.12 (bottom left). These variances within the τB core alone accentuate the distinct DNA binding mechanisms between yeast and human TFIIIC.

Similar to the human structure revealed by cryo-EM, the yeast τB core also supports the winged helix (WH) domains from the N-terminus (1-668) of the τ138 subunit. This includes a total of 4 WH domains and 1 extended WH (eWH) domain, similar to the human τB structure (refer to figure 2.15 for details). The specific domains are WH1 (aa 1-97), WH2 (aa 108-172), WH3 (aa 185-262), WH4 (aa 334-416), and eWH (aa 550-640). Contrary to prior predictions based on sequence analysis [52], this structure reveals that first domain of the subunit τ138 is a WH domain, not a High Mobility Group (HMG) domain as previously thought in yeast. A notable difference between the human and yeast TFIIIC is the support mechanism for WH1; in humans, WH1 is supported by two Zn finger clusters of the subunit TFIIIC90 (see figure 2.14 C), whereas in yeast, the novel form α/β fold from τ60 provides this support. While the yeast TFIIIC complex without DNA has not been solved, based on the similarities between yeast and human τB-DNA subcomplexes, it is hypothesized that the yeast τB subcomplex without DNA might exhibit a similar behavior: with WH1, WH2, and WH4 forming part of the scaffold for initial DNA recognition. This is followed by the interaction with the B-Box upstream region, mediated by the flexible WH3 and eWH domains, a feature observed in the human subcomplex.
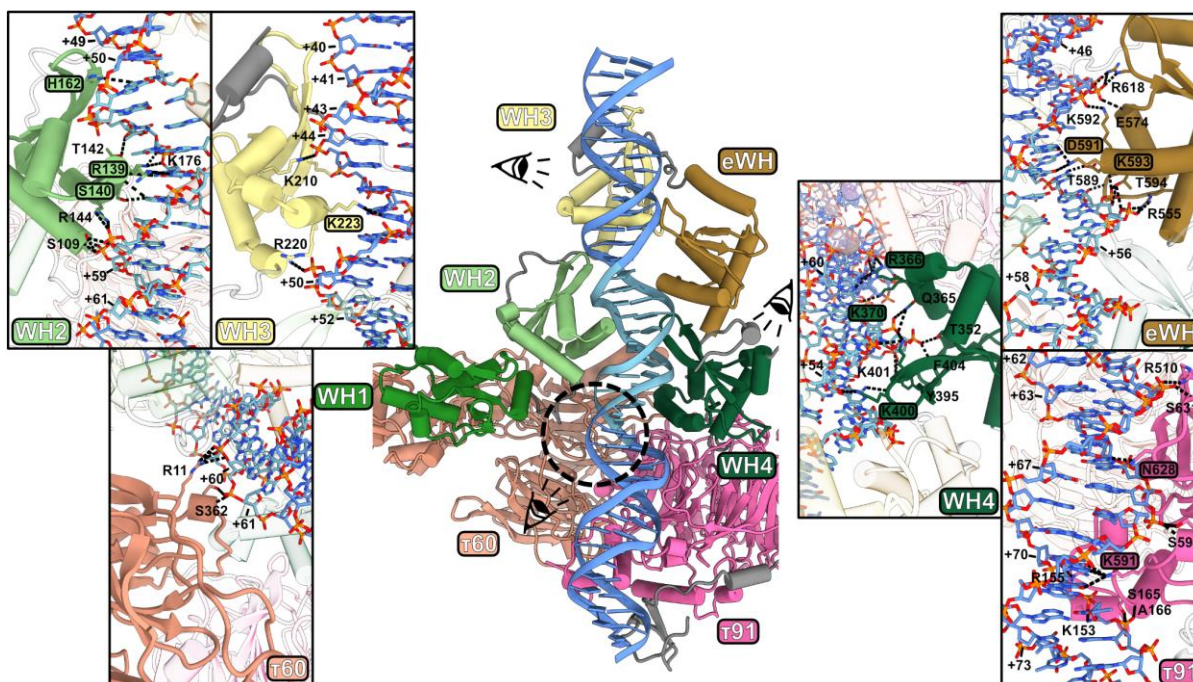
Figure 3.12. Yeast τB subcomplex interaction with the downstream region of the tRNA gene. Cryo-EM structure of the τB-DNA subcomplex. The B-box is highlighted light blue. The interactions between WH2, WH3, WH4 and eWH domains, along with the τ91 and τ60 subunits, with the DNA are detailed in close-up views. Hydrogen bonds are indicated by black dashed lines, with H-bond-forming residues labeled for clarity. Amino acids that interact with the DNA bases are located in colored boxes. Schematic eyes next to the WH3 and WH4 domains and the τ60 subunit demonstrate the perspective from which these inlets are viewed.

The high resolution of the τB-DNA subcomplex enabled the building of 40 base pairs (+37 - +76) of the 3' end of the His_tH(GUG)E2 DNA oligonucleotide, including five additional nucleotides compared to its human counterpart (refer to Figure 2.11). Despite the WH domains and eWH in human and yeast showing similar DNA interactions, distinct structural differences were observed. First, in the human WH1, a long loop (aa 62 to 120) extends and stabilizes WH2 (see Figure 2.14 B). This loop is absent in yeast, making the human WH1 approximately 70 amino acids larger. Second, in yeast, WH2 interacts specifically with the B-Box motif: R139 interacts with C55 on the template strand, G55, and C56 on the non-template strand; S140 also interacts with C56 on the non-template strand; and H162 interacts with G51 on the template strand (refer to Figure 3.12 – top left inlet). Third, the human WH3 domain forms hydrogen bonds with DNA through residues R331, R292, and Y296 (refer to Figure 2.15), whereas in yeast, K210, R220, and K223 are engaged, with K223 uniquely forming base-specific hydrogen bonds with G47 on the non-template strand, a region outside the B-Box. This specificity of K223 is notable, as it represents the only interaction outside the B-Box by any Winged Helix domain in the N-terminus region of τ138. This finding aligns with previous research on other tRNA genes, such as tRNA[tyr] or SUP4, where a mutation at position 45 (G to A, an A nucleotide not typically included in the B-Box) resulted in a fivefold increase in the equilibrium constant of TFIIIC [35]. Our structural data potentially elucidate the significance of this nucleotide's position outside the B-Box, highlighting its role in TFIIIC binding affinity and specificity.

Intriguingly, the region from aa 649 to 668 of τ138 engages with this WH3, forming a small β strand and a helix (see Figure 3.12 – top left region colored grey). This region is part of τIR (aa 641-693), which exhibits high affinity for τ131's TPR domain repeats (aa 123 to 566), when expressed alone (K_d = 80 ± 15 nm) or with a similar affinity when expressed together with eWH in a single construct (K_d = 100 ± 15 nm) measured by Isothermal titration calorimetry (ITC) [115]. It was also found that τIR does not interact with τ131 TPR domains in the cryo-EM structure of TFIIIC involved in type 1 pol III genes, but with τ55 and τ95 [65]. This suggests the versatility of τIR under different conditions. Fourth, in human WH4, R401, R422, and Q423 interact with DNA (see Figure 2.15), whereas in yeast, R366 and K370 interact with G60 on the non-template strand, and K400 interacts with T54 on the template strand and G55 on the non-template strand (refer to Figure 3.12 – middle right). Fifth, the human eWH's K657 interacts with the B-box in a sequence-specific manner, while in yeast, eWH's D591 interacts with C52 on the non-template strand, and K593 interacts with G51 and G52 on the template strand (see Figure 3.12 – top right). These findings, alongside human cryo-EM structures, represent the first detailed exploration of the key amino acids in TFIIIC and crucial nucleotides in a tRNA gene. Prior dimethyl sulfate protection experiments identified vital guanine residues in yeast TFIIIC's interaction with the tRNA3Glu B-box, pinpointing G53 on the template strand and G56, G61, and G62 on the non-template strand (corresponding to G52 on the template strand and G55, G60, and A61 on the non-template strand of the His_tH(GUG)E2 DNA gene) as crucial for the formation of the yeast τB-DNA complex[113].
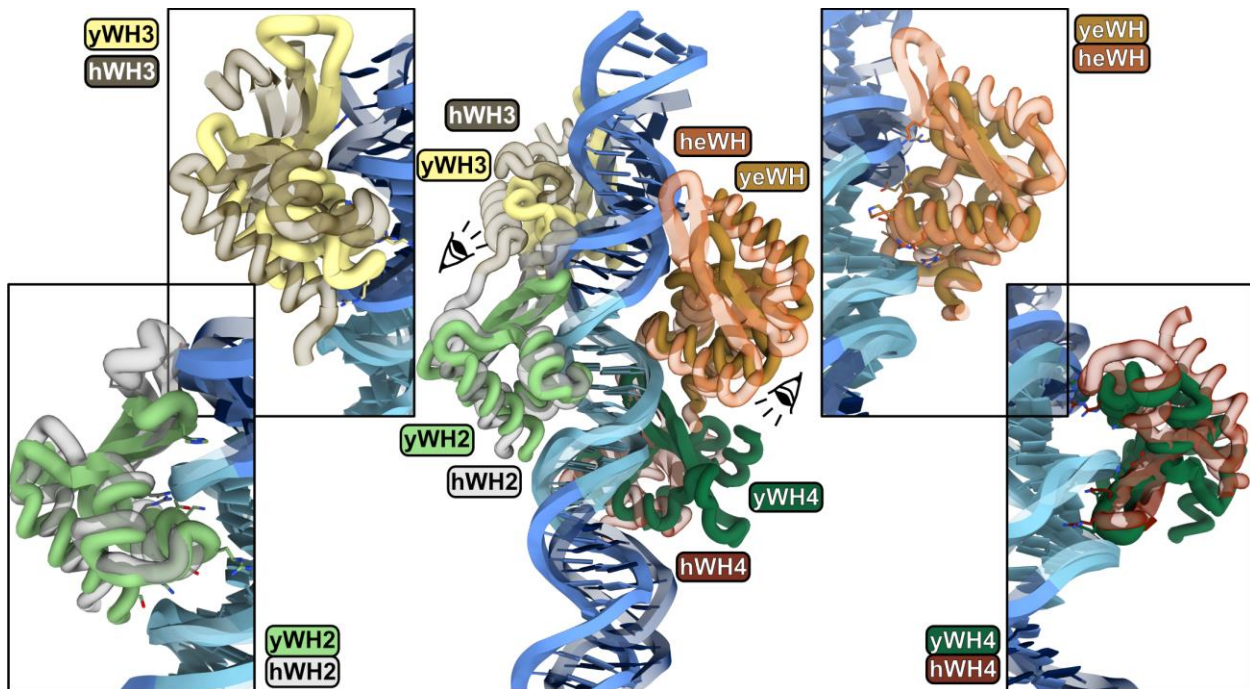
Figure 3.13 Comparison of Winged Helix domains binding in yeast and human τB subcomplexes to tRNA. The WH domains from both yeast and human complexes were superimposed using the B-box promoter region DNA sequence in ChimeraX. Close-up views of each comparison highlight the amino acids responsible for forming hydrogen bonds with the DNA. Schematic eyes positioned next to the WH3 and WH4 domains indicate the viewing angles for these inlets. In the visualization, human WH domains and human tRNA gene are depicted as transparent models to distinguish them from their yeast counterparts.

55

The detailed comparison of yeast and human τB subunits involved in DNA binding, as previously described, is essential for understanding the differences in protein-DNA interactions between these species. However, a broader comparison is beneficial for a more comprehensive perspective. To facilitate this, the cryo-EM structures of the τB subcomplexes bound to DNA from both species were superimposed using the B-Box sequence region (Figure 3.13). Aligning the structures based on the DNA sequence enables an analysis of the spatial arrangement of the various domains in both species. This approach is more objective than selecting a protein for alignment, such as τ138, τ60, or τ91. Notably, despite only 8.1% sequence identity between the yeast and human τ138 subunits[77], the structural configuration surrounding the B-Box, encompassing all WH domains and eWH, is strikingly similar. Furthermore, a comparative analysis reveals that the WH2 and WH4 domains in human and yeast diverge the least in terms of their spatial positioning (Figure 3.13 – bottom inlets), as opposed to WH3 and eWH (Figure 3.13 – top inlets). This observation aligns with the roles of WH2 and WH4 as part of the τB scaffold for initial DNA recognition, anchored to the τB core, which comprises τ60 and τ91 subunits in yeast (discussed in Section 2.6.2).

To gain insights into TFIIIC's role in binding different types of pol III genes, a comparison of τB-DNA complex structures was performed. The remarkable flexibility of the WH3 and eWH domains becomes apparent when examining TFIIIC in the context of type 1 pol III genes in the presence of TFIIIA (Figure 3.14 A). TFIIIA's inclusion significantly alters the interactions of the WH domains and eWH with DNA. In the tRNA gene interaction, all WH domains and eWH engage with the DNA (Figure 3.14 B). However, TFIIIA's presence displaces WH2, WH3, and eWH (Figure 3.12 C – arrows color-coded to match the domains indicate their movement), leaving only WH4 seemingly unaffected by this additional factor. Overall, this comparison shows the pivotal role of TFIIIA in modulating the conformation of TFIIIC, particularly influencing the arrangement of the WH domains and eWH in the τB-DNA complex, and highlights the flexibility of these domains in response to different transcriptional requirements in Pol III genes.
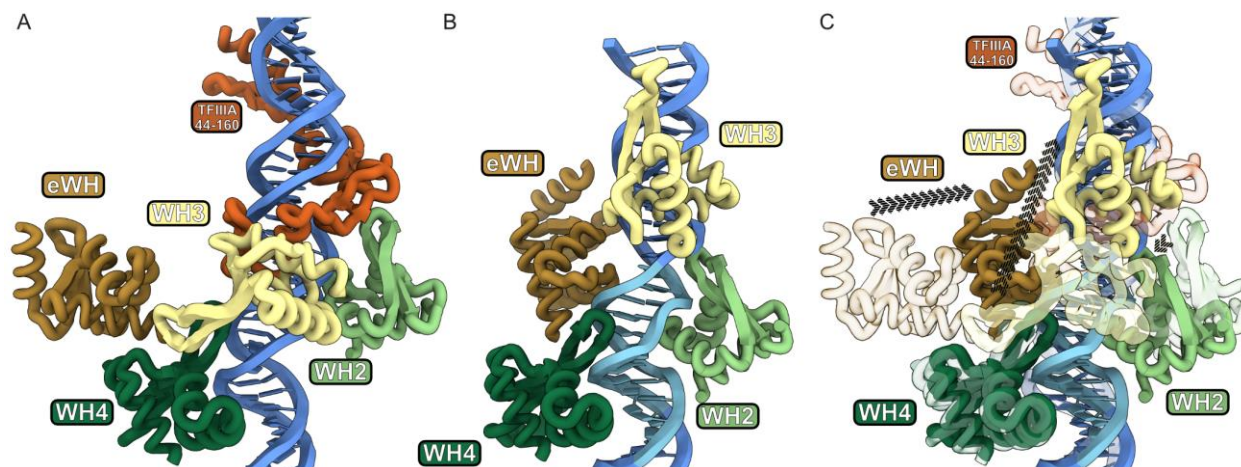
Figure 3.14 Comparison of the yeast τB subcomplexes bound to type 1 and type 2 Pol III genes. A. the τB subcomplex bound to a 5S RNA gene is depicted, where only the Winged Helix (WH) domains are visible and the TFIIIA region from amino acids 44 to 166 is shown, while the rest of the TFIIIA-TFIIIC-TFIIIB complex is excluded (PDB ID: 8FFZ). **B.** Cryo-EM model of the yeast τB subcomplex bound to a tRNA gene, with emphasis on the WH domains. Other domains and loops were hidden for clarity; here, the B-Box promoter is highlighted in light blue. **C.** Superimposition of the τB subcomplexes from Panels A and B using the τ91 and τ60 subunits. In this comparison, the cryo-EM model from the Type 1 τB subcomplex is rendered transparent, and movements of the WH domains are indicated by arrows, color-coded to correspond with each WH domain.

3.3.3 Interaction of τA subcomplex with the tRNA gene

In the human τA subcomplex structure, despite of the presence of its binding region in the TRR-TCT3-2 tRNA gene, a complex with DNA was not resolved. The potential reasons for this are discussed in Section 2.6.3. However, the situation with the yeast TFIIICΔ593-DNA complex is different. A high-resolution structure of yeast τA bound to the tRNA[HIS] gene was successfully resolved at 3.73 Å, as depicted in Figures 3.9 B, C and 3.15 A. In this structure, three domains were identified interacting with the tRNA oligonucleotide: the Helical domain (aa 732 to 838) and WH6 (aa 972 to 1053) from the τ138 subunit, and the C-terminus TPR domain repeats of τ131 (aa 921 to 934), as illustrated in Figure 3.15 B. Analysis of the τA subcomplex's electrostatic (Coulomb) potential surface reveals basic regions within these domains, positioned near the DNA contact points (Figure 3.15 C). Earlier cryo-EM studies of the yTFIIICΔ593-DNA complex highlighted τA's intermittent interaction with DNA (refer to 2D classification in Figure 3.5). With the high-resolution τA-DNA complex now available, this transient binding can be attributed to the Helical domain. Given the current resolution, it is hypothesized that the interactions of these three domains with DNA are non-specific, allowing the τA subcomplex to engage and disengage with DNA while searching for the A-box, potentially becoming stabilized by other transcription factors like TFIIIB [38].
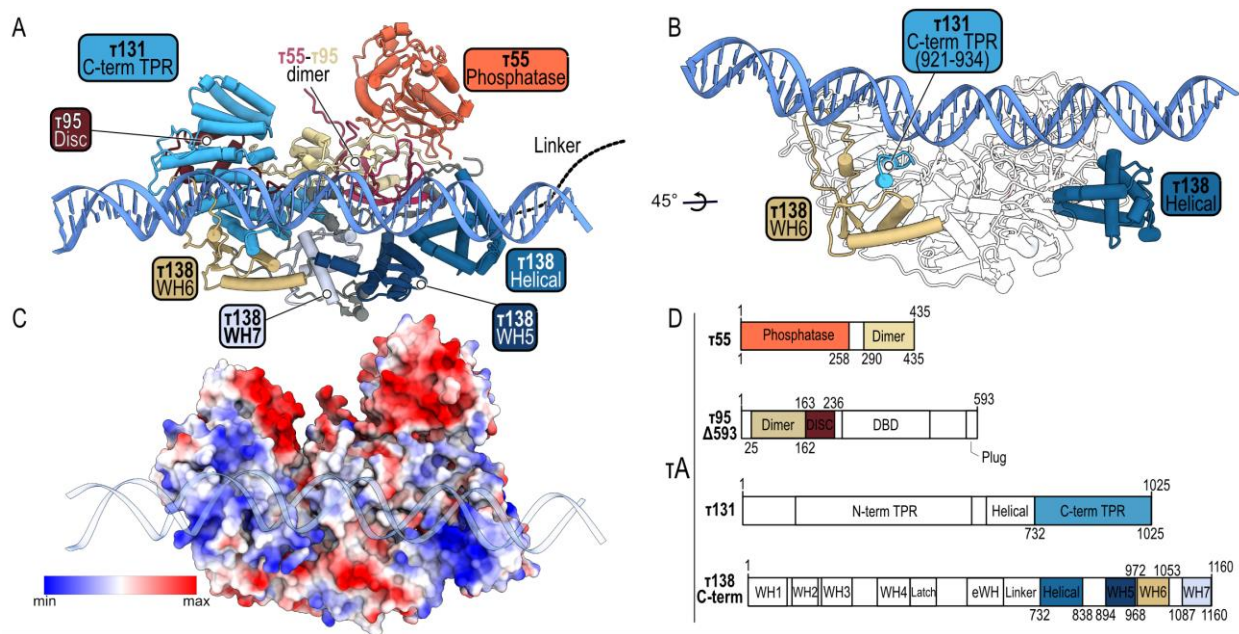
Figure 3.15 Interaction of τA subcomplex with DNA. A. Cryo-EM model of the τA subcomplex, focusing on its domain composition and featuring the flexible linker represented by a dotted line. **B.** Cryo-EM model showing in color only those domains of τA that are actively involved in DNA interaction. **C.** electrostatic potential surface of the τA subcomplex, with the DNA illustrated in a cartoon and transparent style for clarity. **D.** Colored coded domain architecture to match cryo-EM model in panel A. Only domains visible in the structure are colored and the rest are shown in white.

This hypothesis of a 'primarily non-specific τA-DNA interaction' is consistent with earlier studies suggesting a weaker binding affinity between τA and the A-box [113]. Mutational analyses of DNA bases in a tRNA^{tyr} gene have shown that modifications in the B-box more significantly impact TFIIIC binding than alterations in the A-box [35]. Additionally, in efforts to determine the optimal distance between the A- and B-boxes for enhanced binding of yeast TFIIIC to a tRNA gene, two distinct populations of TFIIIC-DNA were identified. This two states appeared when reducing the distance between these motifs by less than 21 nucleotides resulted in protection of both the A-box and an upstream region by τA subcomplex [33]. In yeast RNA polymerase III transcription system reconstitutions, where τA and τB modules were separately reconstituted, the τA subcomplex (this module was composed of τ131, τ95, and τ55 only) exhibited nonspecific DNA binding activity [58]. This is further corroborated by investigations focusing on the DNA Binding Domain in the τA-DNA interaction. For instance, studies on this domain in the Sfc1 subunit of *Schizosaccharomyces pombe* (τ95 in yeast) revealed its ability to bind both double-stranded and single-stranded DNA with micromolar affinity [36]. These findings collectively support the notion of τA's dynamic, primarily non-specific, and relatively weak affinity for DNA, which is crucial for its role in locating its target motif.

The domain architecture of yeast τA when bound to DNA is largely similar to that of the human τA (see domain architecture in Figure 2.11), with some variations. In human τA, the τ138 N-terminus contains 1 homeobox-like domain and 4 Winged Helix domains (WH5 – WH8). Conversely, the yeast τA in complex with DNA comprises 1 Helical domain and only 3 Winged Helix domains (WH5 – WH7), as shown in Figure 3.14 D. For both human and yeast species, this subunit is structurally integral to the formation of the τA subcomplex, primarily due to its interaction with other subunits. In the yeast τA structure, only the C-terminus TPR of τ131 was resolved (see figure 3.15 A and D), in contrast to the human TFIIIC102 where part of the N-terminus was modeled (see figure 2.18 A and C). This difference may correspond to the role of τ131 in the τA-DNA structure, potentially maintaining flexibility to "fish" for the transcription factor TFIIIB, similar to the proposed mechanism in human TFIIIC [66]. In the context of the τ55 subunit in yeast and its human counterpart, TFIIIC35, the dimerization domains are remarkably similar, with the primary distinction being the presence of a phosphatase domain in τ55, which is absent in TFIIIC35. Additionally, TFIIIC63 in humans and τ95 in yeast both exhibit identical dimerization domains and disc domains, further highlighting the structural parallels within these subunits across the two species.

The structural comparison between the yeast τ95 and human TFIIIC63 subunits reveals a striking and unexpected similarity: both structures lack the DNA Binding Domain (DBD), as shown in Figure 2.11 A and Figure 3.15 D. The absence of the DBD in the human unbound τA form could be rationalized, as it might exhibit flexibility while tethered to the TFIIIC63 (τ95 in yeast) subunit, actively seeking its DNA motif, similar to the behavior of WH3 and eWH in τB subcomplex that become ordered only upon DNA binding. The absence of the DNA Binding Domain (DBD) in the yeast τA, even when it is in complex with DNA, emerged as an unexpected observation. This finding was particularly striking given that in a previously resolved yeast τA structure, comprising only the subunits τ131, τ95, and τ55, the DBD was distinctly visible [37]. A plausible explanation for this discrepancy may lie in the missing C-terminus of the τ138 subunit in the earlier structure. This absence could account for the noticeable differences in the structure of that specific yeast τA and the τA-DNA subcomplex examined in this study. To illustrate this point, a comparative analysis with the human τA reveals that the WH6 domain (also present in the τA subcomplex solved in this study) appears to displace the DBD, occupying the same spatial region (as seen in Figure 3.16). A similar displacement is observed in the newly resolved τA-DNA complex. This comparison underscores the significance of analyzing the complete τA structure, including the τ138 C-terminus, to fully understand its function as part of the TFIIIC complex.
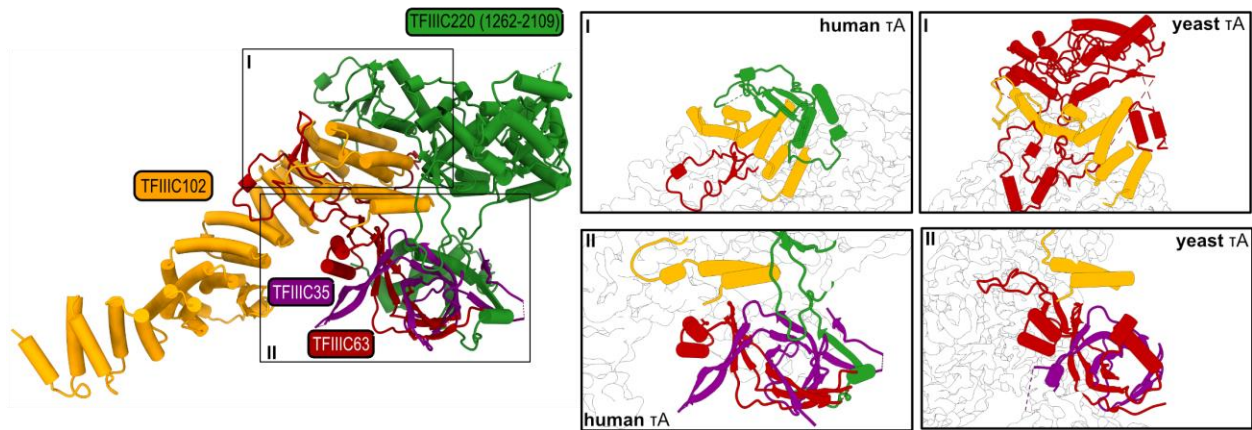
Figure 3.16. A comparison of the human and yeast τA unbound to DNA. An overview of the interactions among various components of the TFIIIC complex, specifically between TFIIIC102, TFIIIC63, and TFIIIC35 subunits, and the C-terminal portion of TFIIIC220, is presented. In humans, the C-terminus TPR domain of TFIIIC102 engages with the WH6 domain of TFIIIC220 (left panel). Conversely, in yeast, the interaction involves the τ131 C-terminus TPR domain and the DNA-binding domain (DBD) of τ95 (right panel). Furthermore, in humans, there is a distinct interaction where the WH8 domain and the internal repeat (IR) of TFIIIC35 within TFIIIC220 associate with the TFIIIC63-TFIIIC35 dimer (left panel). This specific interaction is not observed in the yeast τA subcomplex (right panel) (PDB ID: 6YJ6).

Human TFIIIC220, with 2109 amino acids (see figure 2.11 A), is much larger than the yeast τ138 with 1161 amino acids (see figure 3.9 A). Nevertheless, τ138 contains a similar Number of WH domains, arranged in the same order as in human TFIIIC220, indicating a conserved structure within TFIIIC across these species. The linker segment connecting τ138's N and C termini in yeast, spanning about 64 amino acids, contrasts sharply with its human counterpart of approximately 550 amino acids. Despite the size of this linker is almost 10 times smaller is presumed to be flexible enough to bridge the yeast τA and τB subcomplexes, facilitating simultaneous binding to A- and B-boxes on yeast tRNA genes.

In Section 3.3.2, a comprehensive comparison was made between the τB subcomplexes bound to type 1 and type 2 pol III genes. This analysis revealed a significant rearrangement of domains around the DNA in the presence of the TFIIIA factor (Figure 3.14C). This finding raises the question of whether the τA subcomplexes, when bound to these types of pol III genes, also demonstrate distinct structural configurations. In the context of a type 2 pol III gene, the τA subcomplex engages the tRNA gene via two domains from the C-terminus of τ138 (WH6 and Helical) and one domain from τ131 (Figure 3.15 B). Specifically, the tRNA[His] DNA corresponding traverses the τA subcomplex, crossing the Helical domain at its most 3' end and extending straight to τ138's WH6 domain (Figure 3.17 A). In contrast, when τA is complexed with a 5S rRNA gene in the presence of TFIIIA, the DNA adopts a perpendicular orientation relative to the DNA in type 2 pol III genes (Figure 3.17 B), as elaborated in the close-up views (Figure 3.17C).

Two main factors account for this variation in DNA binding. Firstly, TFIIIA displaces the Helical domain from its DNA interaction site and takes its place (Figure 3.17 C – inlet I). Secondly, the interaction of TFIIIA with τ95's N-terminus region, the acidic tail, known for its role in τA subcomplex autoinhibition

[37] and DBD of Sfc1 (τ95 in yeast) [36], leads to the acidic tail binding to the WH6 domain, essential for tRNA DNA binding (Figure 3.17 C – inlet II). The introduction of TFIIIA thus not only modifies the arrangement of Winged Helix domains in τB but also significantly alters how τA interacts with its target DNA.
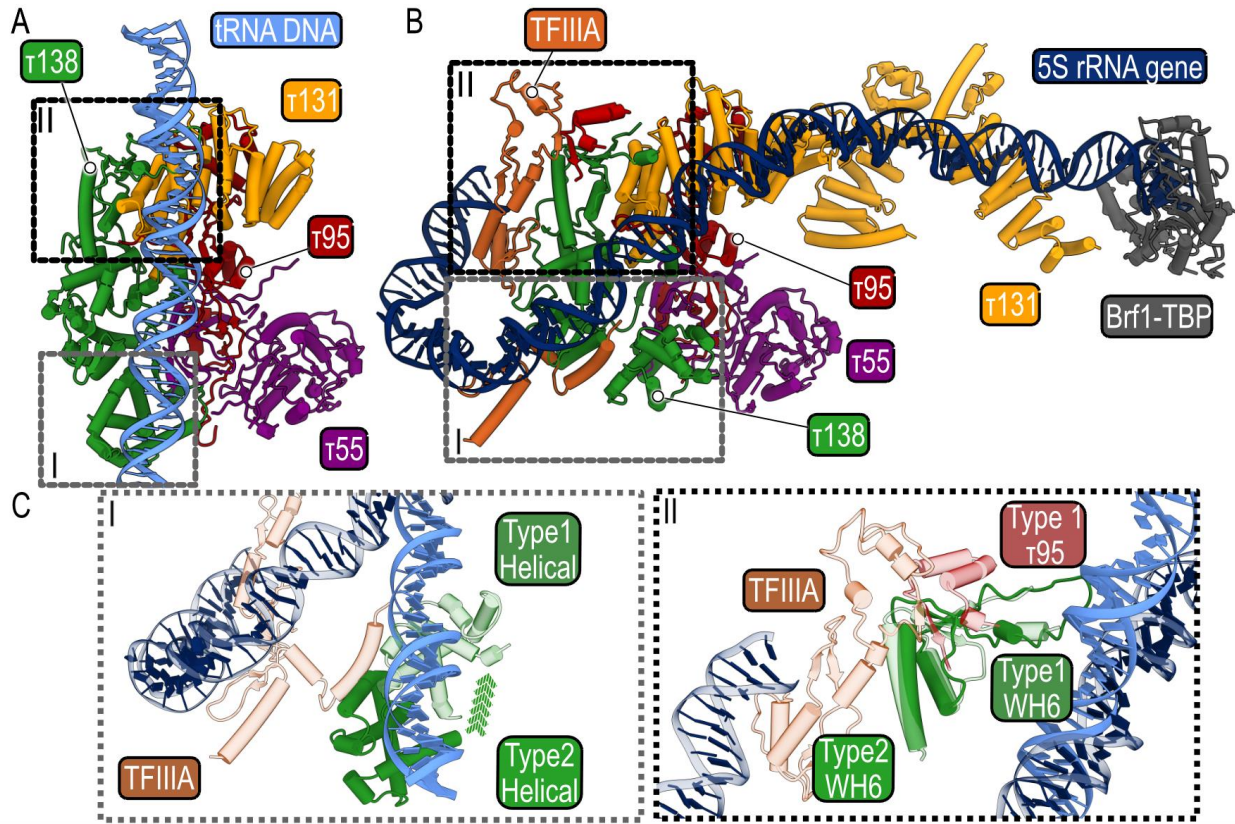
Figure 3.17 Comparison of yeast τA subcomplex bound to type 1 and type 2 pol III genes. A. Cryo-EM model of the τA subcomplex bound to type 2 pol III gene (tRNA). **B.** Cryo-EM model of the τA subcomplex bound to type 1 pol III gene (5S rRNA), TFIIIA (only aa 161 to 509) and Brf1-TBP are included. **C.** Detailed close-up views are provided to highlight the differences, particularly focusing on the movement of the helical domain in τ138 (Inlet I) and the interaction with the WH6 domain in τ138, showing the distinct conformations of the τA subcomplex when bound to these different types of Pol III genes.

Previous studies have established that TFIIIB enhances the stability of the τA-DNA interaction [38]. In this study, adding TFIIIB to the TFIIIC sample was hypothesized to strengthen τA's interaction with its DNA motif and additionally it could aid elucidating the recruitment process of transcription factors in Pol III transcription. To investigate this, the TFIIIC-DNA-TFIIIB complex was analyzed using Electrophoretic Mobility Shift Assay (EMSA) (Figure 3.18 A), and the TFIIIC-DNA-BTB (Brf1-TBP fusion protein) complex was examined through MP (Figure 3.18 B). Despite clear indications of complex formation in these experiments, particularly in the TFIIIC-DNA-TFIIIB complex, where two datasets were obtained, and in the TFIIIC-DNA-BTB complex, for which one dataset was collected (detailed in Table 4.25), no additional densities were detected to signify the presence of these factors in any map produced.
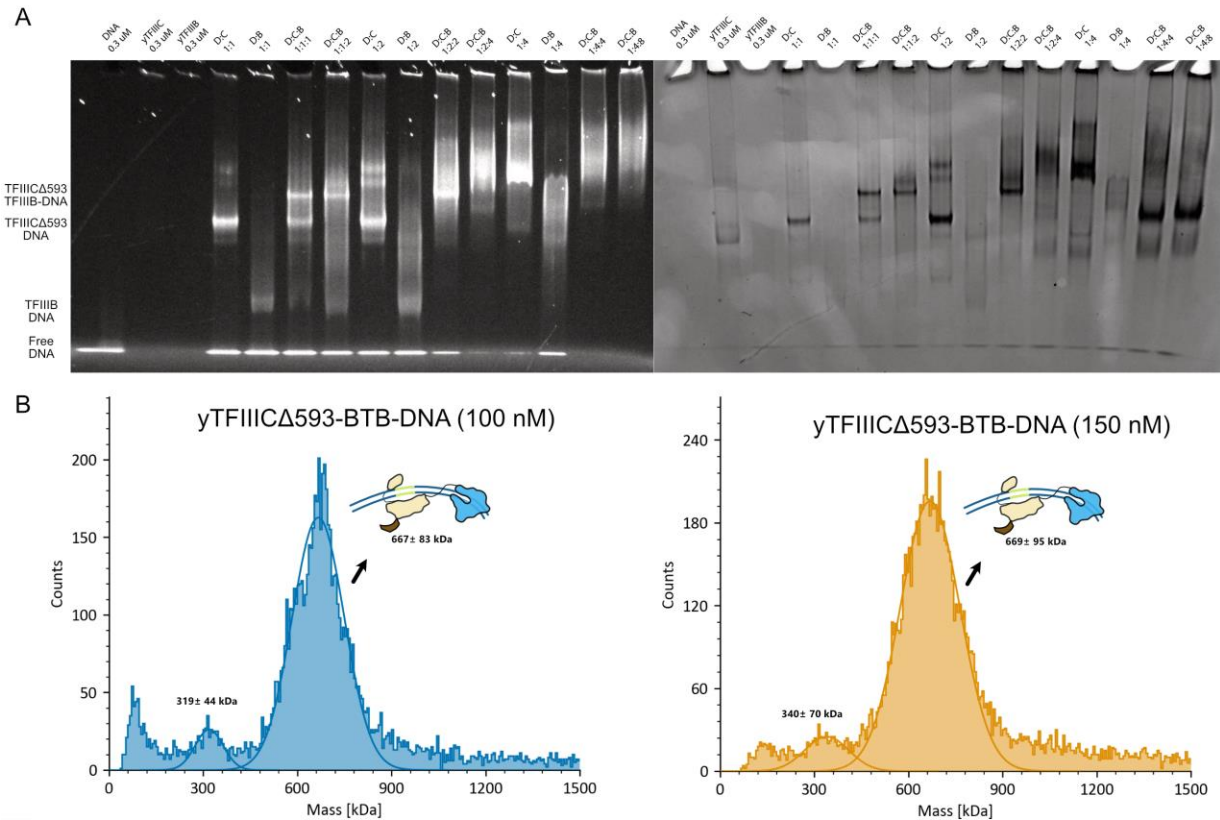
Figure 3.18. Interaction of yTFIIICΔ593 with TFIIIB and BTB (Brf1-TBP fusion protein). A. Electrophoretic mobility shift assay (EMSA) displays the migration patterns of the yTFIIICΔ593-DNA complex when incubated with varying ratios of TFIIIB. 'D' denotes DNA, 'C' represents yTFIIICΔ593, and 'B' indicates TFIIIB. The ratios of each component are labeled above each lane. The left gel was visualized using UV light for DNA detection, while the right gel shows post-staining with Coomassie blue. Higher molecular weight bands are observed when the protein ratio is doubled compared to DNA that might correspond high oligomer states. **B.** Mass photometry histograms demonstrate the formation of stable complexes. On the left, 100 nM yTFIIICΔ593 and 100 mM DNA were incubated with 400 nM BTB, revealing a stable complex of approximately 667 kDa ± 83 kDa (consisting of 520 kDa yTFIIICΔ593, 60 kDa DNA, and 90 kDa BTB). On the right, 150 nM yTFIIICΔ593 and 150 mM DNA combined with 600 nM BTB result in a complex of approximately 669 kDa ± 95 kDa.

Based on the data from the yeast TFIIICΔ593 study, I present an updated model for TFIIIC promoter recognition, incorporating structural insights from human τA unbound to DNA, yeast τA bound to DNA, and both human and yeast τB bound to DNA. Although the focus of this study was not on yTFIIICΔ593 without DNA, the resemblance of its structure to the human τB subcomplex suggests similar functional dynamics in yeast. In this revised model, τB acts as an anchor to the DNA, assisting τA in locating the A-box motif. This search by τA is characterized by dynamic interactions, involving more than just the DNA Binding Domain (DBD) of τ95, and includes a 'fly-casting' mechanism, similar for the subunit TFIIIC63 in humans. The DBD, which is flexibly tethered to τ95, actively participates in this search for the A-box. Consequently, τA's engagement with DNA occurs through two simultaneous activities: the entire complex repeatedly binds and releases from the DNA, while the DBD independently focuses on the A-

box sequence. The binding event might lead to a more ordered structure of the DBD, as illustrated in Figure 3.19.
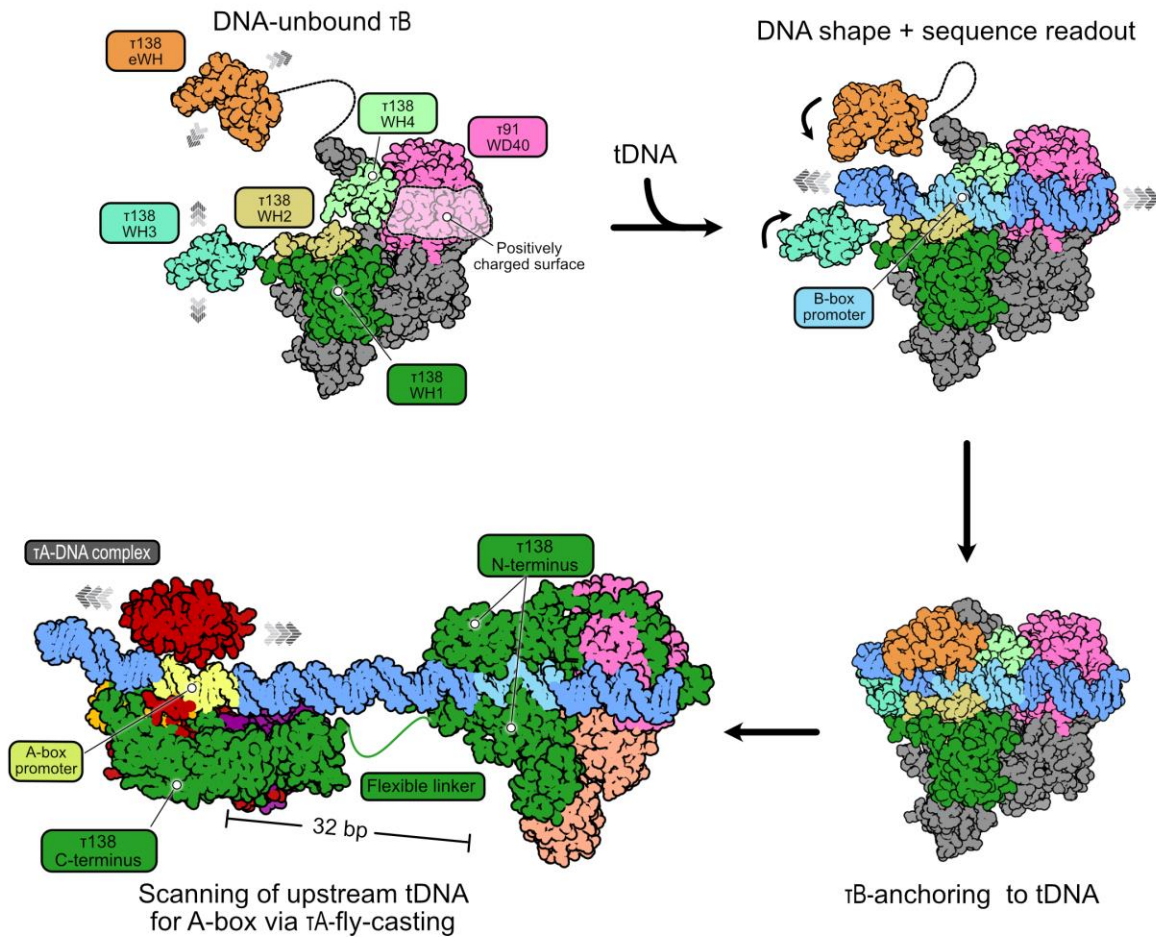
Figure 3.19 Updated model of TFIIIC promoter recognition. DNA recognition starts with the shape and sequence readout of the DNA, followed by the anchoring of TFIIIC to the DNA through its τB subcomplex. The τA subcomplex, connected to τB via a flexible linker, actively searches for its A-Box motif. This search is not limited to the DNA binding domain of τA but also involves other domains, enabling the τA subcomplex to bind and unbind from the DNA until it locates its specific motif. The model suggests that this highly dynamic interaction of τA with the DNA may be stabilized upon the association of the Brf1-TBP subunits, hinting at a cooperative mechanism in the promoter recognition process by TFIIIC.

3.4 Single molecule fluorescence microscopy applied to yTFIIIC

Despite obtaining a high-resolution structure of the yeast τA bound to DNA, some aspects remained elusive, attributed to the subcomplex's dynamic nature. The initial cryo-EM analysis of the first dataset revealed a fuzzy region in the 2D classes and a unidentified density at the top of the τA subcomplex obtained by a 3D variability analysis with cryoSPARC (as shown in Figure 3.5). This region, which was later speculated to be the DBD connected to the τ95 subunit within the τA-DNA subcomplex, could not be distinctly resolved in the obtained maps. By lusing all the available datasets, a more detailed representation of various potential states of the τA subcomplex at the 2D level was achieved (refer to Figure 3.20 A and B). These 2D classes not only show the τA subcomplex's dynamic binding behavior with DNA but also consistently exhibited the previously mentioned fuzzy region.
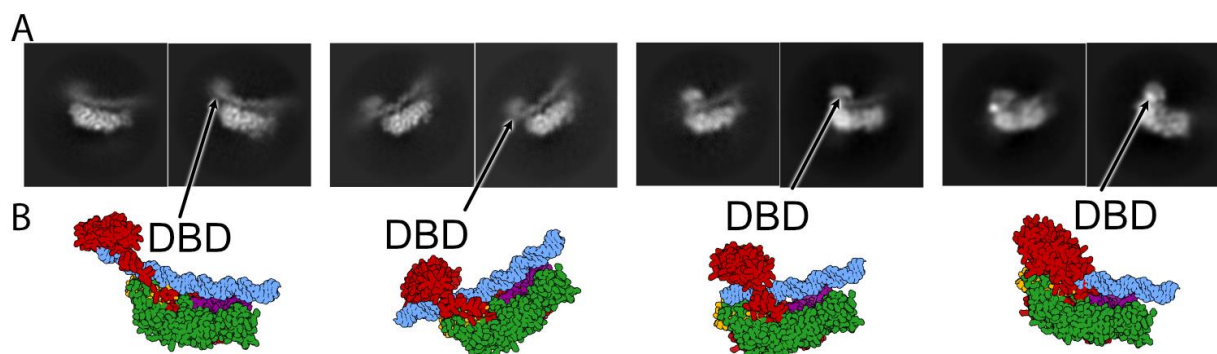
Figure 3.20 Potential dynamic binding of the DNA binding domain (DBD) of the τA subcomplex. A. Representative 2D classes from all the datasets showing the dynamic movement of the DBD (τ95). B. Schematic representations showing the movement of the DBD in the representative 2D classes.

Considering these observations, an alternative approach was used to investigate the role of the DBD in recognizing the A-box, a question that remained open after the cryo-EM analysis. This method involved single molecule fluorescence microscopy. The strategy consists in tagging the yTFIIICΔ593 complex with ybbR near the DBD. This tag is then utilized for the specific addition of a dye, such as Cy5 or Cy5.5, facilitated by an Sfp enzyme. Additionally, a donor dye needed to be incorporated into the tRNA, particularly near the A-box, for Förster Resonance Energy Transfer (FRET) analysis between the protein and DNA. To achieve this, modifications were made to the His_tH(GUG)E2 gene, as detailed in Section 4.7.2.

The preparation of this modified template involved a two-step PCR process. The initial phase was an analytical PCR gradient, conducted to ascertain the optimal temperature for maximizing oligonucleotide production. As indicated in Figure 3.21 A, the most effective yield was obtained at 59.7 °C. Subsequently, this temperature setting was applied in the second phase, a preparative PCR, which aimed to produce an adequate amount of the template for the planned TIRF experiments.
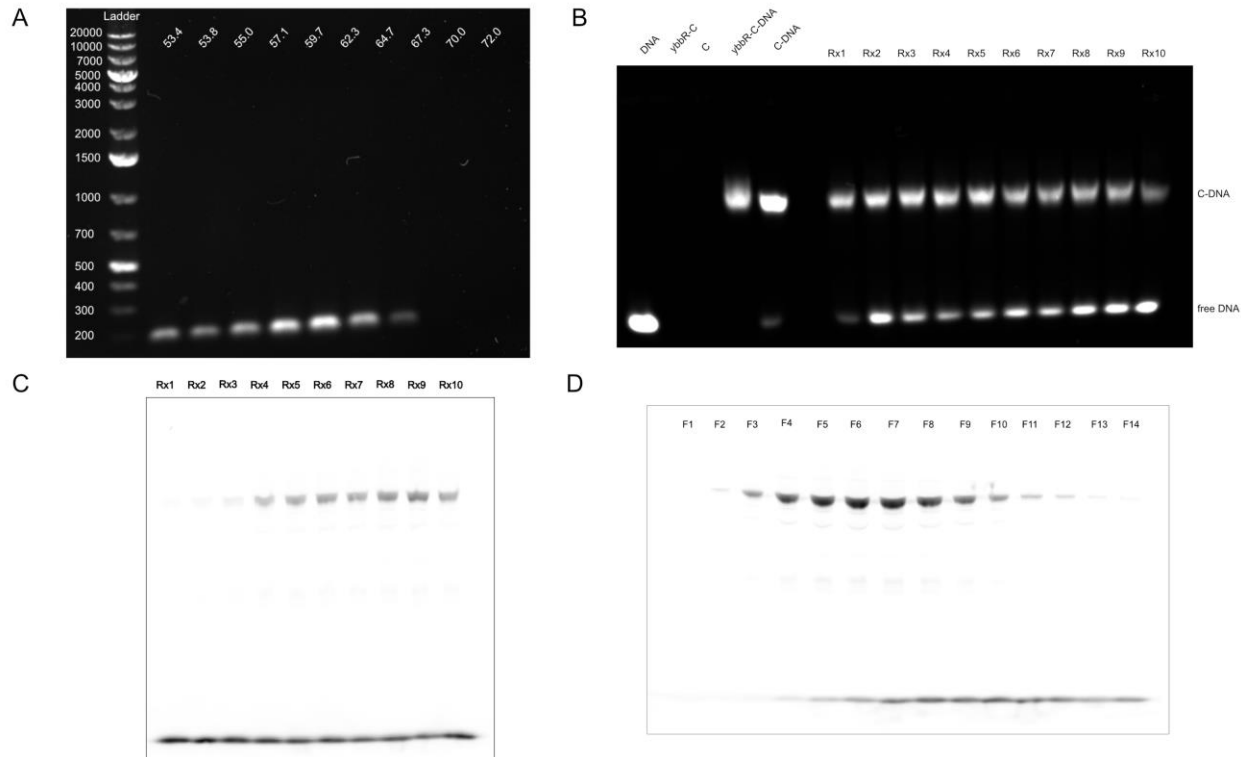
Figure 3.21 Preparation of ybbR-yTFIIICΔ593 for TIRF Experiments. A. An agarose gel showing the results of an analytical PCR gradient to add adaptors to the 3' and 5' region of the His_tH(GUG)E2 gene for TIRF experiments. **B.** Electrophoretic Mobility Shift Assay (EMSA) assessing the functionality of Cy5-labeled ybbR-yTFIIICΔ593. 'C' denotes the yTFIIIC complex, and 'ybbR-C' indicates the ybbR-yTFIIIC complex. Ten different reactions were conducted, varying the incubation time and concentration of the Sfp enzyme. **C.** SDS-PAGE analyzed with a Typhoon instrument in the Cy5 channel, displaying the optimization process for labeling ybbR-yTFIIICΔ593. This step was crucial to determine the optimal conditions for labeling without compromising the complex's integrity. **D.** SDS-PAGE imaged with the Typhoon, presenting all fractions collected after Protein Desalting (PD)-column purification following preparative labeling. The best condition was applied to the entire batch of ybbR-yTFIIICΔ593 for subsequent experiments.

Recognizing the dynamic nature of the τA-DNA subcomplex, Simone Höfler suggested introducing a ybbR tag at positions 286 (after) and 506 (before) the DNA Binding Domain (DBD) in the τ95Δ593 subunit to study its interaction with DNA. This approach was informed by AlphaFold-Multimer predictions[123] and the experimental structure of the τA-DNA subcomplex obtained in this project. Despite several cloning attempts, initial efforts to add this tag were unsuccessful. Troubleshooting strategies, including the use of an in-house Phusion master mix and the Platinum Hot Start PCR Master Mix (ThermoFischer), as well as adjusting PCR settings with different annealing temperatures and amplification cycles, were employed without success. However, Jonas Weidenhaussen, utilizing a Gibson assembly approach, successfully added the tag to τ95 at positions 286 and 506.

Subsequently, I proceeded with the Gibson assembly of all yTFIIIC subunits into a single vector. Only the variant with the tag at position 286 was successfully assembled, and this variant was chosen for further work. Standard baculovirus production and purification methods, as described in Section 4.6.3, were

applied to the new ybbR-yTFIIICΔ593 complex. Additionally, this batch underwent a labeling optimization process (outlined in Section 4.7.1). This step included assessing the complex's integrity (Figure 3.21 B) and determining optimal labeling conditions (Figure 3.21 C). Based on these results, the conditions from reaction 7 (Rx 7) were selected for labeling the entire batch of freshly purified ybbR-yTFIIICΔ593 complex. Finally, to remove unincorporated dye, the complex underwent purification using two consecutive Protein Desalting (PD) columns. The fractions were evaluated by SDS-PAGE (see figure 3.21 D) and those fractions with minimal or no free dye were pooled, aliquoted, and flash-frozen for future use. These steps facilitated the production of ybbR-yTFIIICΔ593-Cy5 and ybbR-yTFIIICΔ593-Cy5.5, which were utilized in TIRF experiments conducted by Anastasiia Chaban (refer to Section 4.7.3). A preliminary experiment was performed and a representative TIRF trace is analyzed (Figure 3.22). This trace reveals that the interaction of yTFIIIC with the tRNA gene is dynamic. While it was hypothesized that the τB subcomplex anchors TFIIIC to the DNA, enabling the τA subcomplex searching for the A-box, this trace suggests that multiple yTFIIIC complexes bind sequentially to the same tRNA gene. Further experiments are necessary to corroborate this observation.
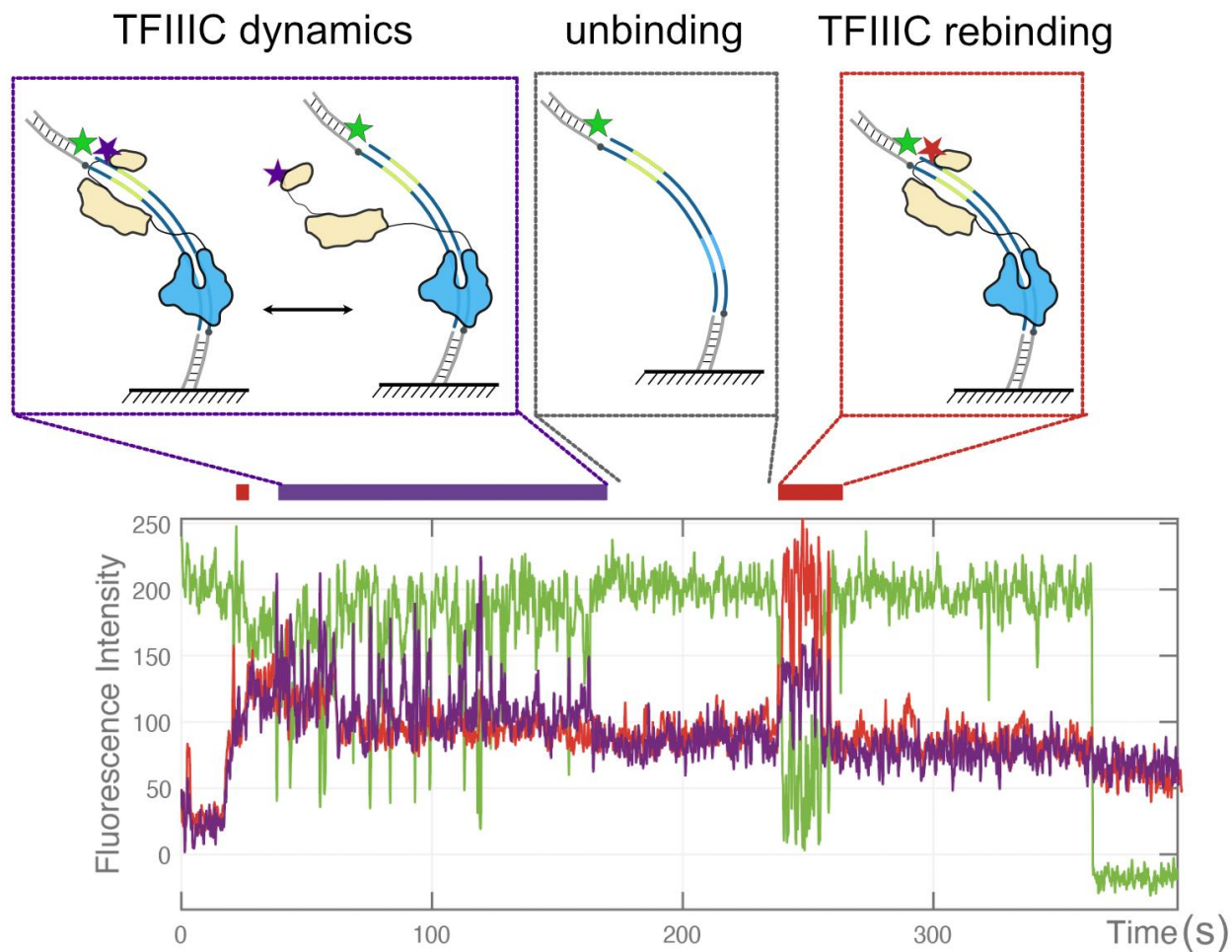
Figure 3.22 Representative TIRF microscopy trace of ybbR-yTFIIICΔ593-Cy5 and ybbR-yTFIIICΔ593-Cy5.5 interacting with a tRNA gene. This trace illustrates the binding behavior of two distinct yTFIIICΔ593 complexes over a period of 10 min. The yTFIIICΔ593 complex labeled with Cy5.5 demonstrates a prolonged and dynamic binding event, lasting approximately 2 minutes. This is followed by a period with no binding. Subsequently, a shorter binding event is observed from the yTFIIICΔ593 complex labeled with Cy5.

3.5 Future perspectives

Looking ahead, the elucidation of human TFIIIC structures in my study, both with and without bound DNA, has set a new precedent in understanding the promoter recognition at resolutions between 3.2 to 3.5 Å. These high-resolution maps have revealed the subunit composition of τA and τB subcomplexes, unmasking the pivotal role of TFIIIC220 in these complexes. The discovery of the flexible linker's location, instrumental for TFIIIC's binding to A- and B-box promoters, emerged from a comprehensive analysis of cryo-EM data. This data included precise measurements of the spatial relationships between τA subcomplex and τB-dimer particles. The proposed mechanistic model of TFIIIC's interaction with tRNA genes, particularly the role of τB in anchoring the complex to DNA to facilitate τA subcomplex's searching for its promoter through a fly-casting mechanism, improve our understanding from a structural point of view

Furthermore, the resolution of yeast TFIIIC structures fully bound to DNA at 2.67 Å for the τB–DNA complex and 3.73 Å for the τA-DNA subcomplex marks a significant advancement. These structures not only affirm a promoter recognition mechanism similar to that in humans but also introduce novel insights. The τB-DNA complex illuminates aspects of promoter recognition, mirroring the role of TFIIIC220 in humans, as seen in τ138. The τA-DNA complex offers a first-time glimpse into this dynamic interaction, further substantiated by single-molecule fluorescent microscopy (SMFM) experiments.

The future of elucidating the dynamic interactions of the τA subcomplex, particularly the role of the DNA Binding Domain, appears promising, especially with the integration of Single-Molecule Fluorescence Microscopy (SMFM) alongside cryo-Electron Microscopy (cryo-EM). While the high-resolution structure of the τA subcomplex has been determined, unanswered questions persist, emphasizing the need for employing SMFM to clarify its function in promoter recognition. Moreover, understanding how TFIIIC recruits TFIIIB remains a complex challenge. Despite successful demonstrations of complex stability using Electrophoretic Mobility Shift Assays (EMSA) and Mass Photometry (MP), difficulties arise in maintaining this stability during cryo-EM sample preparation. This suggests that the interaction between TFIIIB and TFIIIC, whether partial or complete, might be more dynamic than can be effectively captured by cryo-EM. This scenario underscores the necessity of exploring additional methods or developing innovative strategies to visualize and comprehend these dynamic interactions in the context of transcription initiation.

III. Material and Methods

4. Material and methods

4.1 DNA oligonucleotides preparation

DNA oligonucleotides corresponding to the human TRR-TCT3-2 gene and the yeast His_tH(GUG)E2 gene were ordered from Sigma-Aldrich. These oligonucleotides were utilized for experiments involving cryo-EM, Mass Photometry, and EMSA. The annealing process for both nontemplate and template strands was conducted in dH_2O, beginning with heating to 95°C for 5 minutes and followed by a controlled cooling to 20°C at a rate of 1°C per min. After annealing, the DNA underwent size-exclusion chromatography on a Superdex 200 Increase 3.2/300 (Cytiva) column, equilibrated with DNA buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM MgCl2, and 5 mM DTT). This chromatography step was crucial for removing any incomplete annealing products and impurities. Subsequently, the major peak fractions were collected, and their concentration was accurately quantified using a Nanodrop spectrophotometer.

4.2 Electrophoretic mobility shift assay

4.2.1 Human TFIIIC-DNA complex

To test the interaction of TFIIIC and DNA, I performed an EMSA experiment. An amount of 200 nM of double-stranded human TRR-TCT3-2 DNA oligonucleotides were used, mixed with varying concentrations of the hTFIIIC (100, 200, 300, 400, and 800 nM) in a buffer comprising 20 mM HEPES pH 8, 225 mM KCl, 2 mM MgCl2, 5 mM DTT. This mixture was incubated at RT for 10 min. To each of these mixtures, glycerol diluted in the previous buffer was added to achieve a final concentration of 10% as the loading buffer. These preparations were then loaded onto a 0.8% agarose gel made with 0.5× tris-borate buffer (44.6 mM Tris, 44.5 mM boric acid, pH 8.3) and with 3 µL of SYBR Safe DNA stain (Life Technologies). Electrophoresis was conducted at 85 V for 45 min in a 4°C cold room, using the same buffer as the agarose gel. After the electrophoresis was done, the gels were imaged using the Quantum CX5 imaging system from Vilber.

4.2.2 Yeast TFIIIC-DNA-TFIIIB-fpt1 complex

To evaluate the interactions of fpt1 with both TFIIIC-DNA complex and TFIIIC-DNA-TFIIIB complex, Alicia Santos-Aledo, under my supervision, performed this assay using two separate reaction setups. The DNA used was the His_tH(GUG)E2 gene fragment (-43 to +76), which includes the upstream region required for TFIIIB binding. The first part of this protocol involved mixing TFIIICΔ593 (300 nM), DNA (300 nM), and fpt1 protein (provided by Marlize van Breugel) at varying concentrations (300 nM, 600 nM, 1.2 µM, 2.4 µM) and incubating for 10 min at RT in a buffer composed of 20 mM HEPES pH 8, 150 mM KCl, 2 mM MgCl2, and 5 mM DTT. In the second mixture, after an initial incubation of TFIIIC and DNA (both at 300 nM), brf1-TBP and BDP1, each at 300 nM, were added sequentially to the mixture, with 5 min of incubation after each addition. Fpt1 was then added to this complex at varying concentrations (300 nM, 600 nM, 1.2 µM, 2.4 µM) and incubated for a further 10 min. Following incubation, 10% glycerol (diluted in the same buffer) was added to each mixture as a loading buffer. For the electrophoresis, a 3-8%

precast polyacrylamide gel (ThermoFisher) was used. The running buffer for the gel was a Tris/glycine buffer, which was prepared by dissolving 3.3 g of Tris, 14.4 g of glycine and 154 mg of DTT in 1 L of ddH$_2$O. Electrophoresis was carried out at 85 V for 45 min at 4°C. Afterwards, the gel was incubated with 1:10000 dilution of SYBR Gold Nucleic acid gold stain (Invitrogen) for 10 min and then visualized using a Quantum CX5 imaging system from Vilber.

4.3 Filter binding assay

Before initiating the filter binding assay to evaluate the binding affinity of human TFIIIC for its DNA target, Florence Baudin carried out the radioactive labeling of both the template and nontemplate oligonucleotides of the human TRR-TCT3-2 gene (-8 to +91). The 5' ends of these oligonucleotides were radioactively marked with [γ-32P] adenosine 5'-triphosphate, utilizing T4 polynucleotide kinase. The labeled oligonucleotides were separated on a denaturing gel composed of 10% acrylamide/bisacrylamide with 8.3 M urea. Following electrophoresis, the band specific to the full-length oligonucleotide was cut and the oligonucleotides were eluted overnight in a buffer containing 0.5 M ammonium acetate, 10 mM magnesium acetate, 0.1% SDS, and 0.1 mM EDTA. The eluted oligonucleotide was subsequently concentrated by ethanol precipitation.

In the filter binding assay performed by Florence Baudin, the radiolabeled oligonucleotides were hybridized to their complementary strands in a solution of 20 mM HEPES pH 7.5, 5 mM MgCl2, and 100 mM KCl, incubated at RT for 30 min. A fixed quantity of the radiolabeled DNA (approximately 0.5 nM) was incubated with increasing concentrations of TFIIIC (ranging from 0.1 nM to 1 µM) in a buffer composed of 20 mM Tris-HCl (pH 7.5), 150 mM KCl, 2 mM MgCl2, and 5 mM DTT at 4°C for 30 min. The resulting DNA-protein complexes were passed through a 0.45-µm nitrocellulose filter, and the filters were then measured for radioactivity using a Tri-Carb 2800TR Cerenkov scintillation counter from PerkinElmer. The recorded counts were normalized and analyzed using a Hill equation with a set Hill coefficient of 1, employing the Prism software by GraphPad for the curve fitting.

4.4 Mass Photometry experiments

Karine Lapouge prepared high-precision microscope coverslips (24 mm x 50 mm) for each mass photometry experiment as follows: The coverslips were sequentially washed with ddH$_2$O and isopropanol, alternated twice, with a final ddH$_2$O rinse. They were then dried using pressurized air. To define reaction areas, a silicone gasket with six cavities was centrally placed on each coverslip.

For the hTFIIIC complex, prior to measurements, 19 µL of specific buffers were applied to these holes for autofocus stabilization: MP Buffer 1 (20 mM HEPES pH 8, 150 mM KCl, 2 mM MgCl$_2$, 5 mM DTT) for hTFIIIC samples, and MP Buffer 2 (20 mM HEPES pH 8, 200 mM KCl, 2 mM MgCl$_2$, 5 mM DTT) for hTFIIIC-DNA samples. An 1 µL aliquot of the sample at 400 nM concentration was then introduced into the buffer.

To evaluate the effect of salt concentration on the binding stability of the yTFIIICΔ593 complex with the 85 bp and 120 bp His_tH(GUG)E2 gene. A buffer composed of 20 mM HEPES pH 8, 2 mM MgCl$_2$, 5 mM DTT, with KCl concentration ranging from 150 mM to 250 mM (in steps of 25 mM) were used. For the measurements, 19 µL of the respective buffers were first dispensed into the designated wells to

facilitate autofocus stabilization. Subsequently, a 1 μL aliquot of the yTFIIICΔ593 complex, at a concentration of 400 nM, was introduced into the buffer for analysis.

To assess the concentration-dependent oligomerization of yTFIIICΔ593 and yTFIIICΔ593-DNA complexes, a buffer containing 20 mM HEPES pH 8, 100 mM KCl, 2 mM MgCl$_2$, and 5 mM DTT was used. To achieve varying concentrations, the experimental setup was adjusted. Instead of the standard 1 μL of sample plus 19 μL of buffer, increased volumes of the initial sample were used. Specifically, 2.5 μL of the sample was added to get a final concentration of 50 nM, and 5 μl for a concentration of 100 nM, with the buffer volume adjusted accordingly.

I performed the measurements using a Refeyn TwoMP mass photometer (Refeyn Ltd., Oxford, UK). One-min video recordings were captured with AcquireMP software (Refeyn Ltd., version 2.4.0), setting the image size to 150 × 59 binned px, which translates to a 10.9 μm x 4.3 μm imaged area and a 46.3-μm^2 detection area. Analysis of the data was performed using DiscoverMP software (Refeyn Ltd., version 2.4.0). For the creation of a standard contrast-to-mass calibration curve, which achieved a linear regression fit with an R^2 of 0.99999, proteins like bovine serum albumin (66-kDa) and immunoglobulin G (150- and 300-kDa) were used, diluted in working buffer.

4.5 Cloning

4.5.1 Recombinant human TFIIIC

The coding sequences for the six hTFIIIC genes, codon-optimized for expression in insect cells, were ordered from DNASU. Three of these genes required sequence modifications via PCR to match their protein sequences with those listed in UniProt (see table 4.1).

Table 4.1 Modifications in hTFIIIC genes applied to the original DNASU sequences

Gene	Protein	Modification
GTF3C1	TFIIIC220	No modification
GTF3C2	TFIIIC110	First 12 aminoacids from N-term removed
GTF3C3	TFIIIC102	S70N
GTF3C4	TFIIIC90	No modification
GTF3C5	TFIIIC63	No modification
GTF3C6	TFIIIC35	S1M

To introduce these modifications, I used a strategy similar to that outlined in Step 2 of the restriction-free method, as described below (see table 4.4 for reaction details). To facilitate the deletion of specific regions or introduction of point mutations in a plasmid, a pair of primers was designed. These primers are complementary to the sequences adjacent to the desired modification site, and each was designed with a melting temperature (T$_m$) of 65°C to ensure efficient annealing

Once all hTFIIIC genes were matching with the sequences described in UniProt, the next step involved their insertion into the pACEBac vector (Geneva Biotech). For this, I applied a restriction-free method, a protocol originally developed by Van Den Ent [124] and later adapted in our lab by Helga Grötsch. This

technique involved two successive Polymerase Chain Reactions (PCR). The initial step consisted in using a pair of primers: the initial primer, with a T_m of 57 °C, is designed to anneal the target insert, while the second primer, with a T_m of 64 °C, anneals to the target vector. In this first PCR, the DNASU plasmid containing each hTFIIIC gene served as the template in the PCR mixture (refer to Table 4. 2), using the PCR settings outlined in Table 4.3.

Table 4.2 PCR mix components

Component	Amount
5x High-Fidelity PCR Green buffer (ThermoFisher)	10 µL
dNTPs (10 mM)	2 µL
DNA template	20 ng
Primers 1 and 2 (20 pmol/ µl)	1.25 µL
Phusion polymerase	0.5 µL
Water	Up to 50 µL

Table 4.3 PCR settings. 35 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	30 s
Denaturation	98 °C	30 s
Annealing	55 °C	30 s
Extension	72 °C	20-30 s/kb
Final extension	72 °C	2x extension time

The whole PCR reaction from the preceding step was loaded on a 1% agarose gel, which was prepared with 3 µL of SYBR Safe DNA stain (Life Technologies) in 1X TBE buffer. Electrophoresis was performed at 100 V for 45 min at RT. Subsequently, the band of the expected size was excised and subjected to gel purification using the MinElute Gel Extraction Kit (Qiagen), according to the manufacturer's protocol.

The purified DNA fragment was then used as a megaprimer in the second PCR step. Detailed information regarding the PCR components and settings can be found in Table 4.4 and Table 4.5, respectively.

Table 4.4 PCR mix components to amplify hTFIIIC into pAceBac vector

Component	Amount
5x High-Fidelity PCR Green buffer (ThermoFisher)	10 µL
dNTPs (10 mM)	2 µL
Vector	20 ng
Megaprimer or mutation primer	100 ng
Phusion polymerase	0.5 µL
Water	Up to 50 µL

Table 4.5 PCR settings to amplify hTFIIIC into pAceBac vector. 18 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	30 s
Denaturation	98 °C	30 s
Annealing	55 °C	45 s
Extension	72 °C	1 min/kb, >7kb: 1.5 min/kb
Final extension	72 °C	2x extension time

DpnI Fast Digest enzyme (NEB) was added to the PCR mixture to remove the starting vector. The resultant product was used to transform XL1-Blue chemocompetent cells, which were subsequently spread on agar plates containing the relevant antibiotics. Successful colony formation, prompted the

selection and growth of several colonies in separate tubes containing 3–5 mL of LB medium supplemented with the previously used antibiotic. Plasmid DNA was then isolated from these cultures using a Miniprep kit (ThermoFisher) according to the manufacture's guidelines. To verify the insertion of all sequences into the pACEBac vector, Sanger sequencing was performed, with samples sent to Eurofins Genomics for analysis.

After confirming successful integration of the hTFIIIC genes into the pACEBac vectors, I introduced a DNA sequence coding for tobacco etch virus–cleavable (TEV) His-tag at the N-terminus of GTF3C2 and GTF3C5, before starting the biGBac cloning process. This insertion was done using an overlap extension PCR method, similar to the second PCR step in our restriction-free cloning protocol. Detailed procedures and parameters for this step can be found in Tables 4.4 and 4.5. Additionally, a 3x FLAG-tag, amplified from a donor vector, was inserted at the C-terminus of GTF3C1 utilizing the restriction-free cloning method previously described.

For the assembly of a six-gene TFIIIC expression construct, the biGBac system was used [109], this protocol was further adapted by Mathias Girbig in our lab. The τA subcomplex genes (GTF3C3, GTF3C5 and GTF3C6) and the τB subcomplex genes (GTF3C1, GTF3C2 and GTF3C4) were independently cloned into pBIG1a and pBIG1b vectors, respectively. To achieve the insertion of the hTFIIIC genes into their corresponding biGBac vector the following set of primers were used (see table 4.6).

Table 4.6: Sets of primers used for efficient BiGBac cloning of τA and τB genes into pBIG1a and pBIG1b vectors

biGBac vector	Gene	Primers	Primer sequence
pBIG1a	GTF3C6	CasI forward and CasI reverse	CasI fw: AACGCTCTATGGTCTAAAGATTTAAATCGACCTACTCCGGAATATTAATAGATC
	GTF3C5	CasII forward and CasII reverse	CasII fw: AAACTGGATACTATTGCACGTTTAAATCGACCTACTCCGGAATATTAATAGATC
	GTF3C3	CasIII forward and CasV reverse	CasIII fw: AAACCTAATGATGCCTGATGTTTAAATCGACCTACTCCGGAATATTAATAGATC
pBIG1b	GTF3C1	CasI forward and CasI reverse	CasI rv: AAACGTGCAATAGTATCCAGTTTATTTAAATGGTTATGATAGTTATTGCTCAGCG
	GTF3C2	CasII forward and CasII reverse	CasII_rv: AAACATCAGGCATCATTAGGTTTATTTAAATGGTTATGATAGTTATTGCTCAGCG
	GTF3C4	CasIII forward and CasV reverse	Casw_rev: AACCCCGATTGAGATATAGATTTATTTAAATGGTTATGATAGTTATTGCTCAGCG

The τA and τB genes were then amplified using the following PCR mix components (see Table 4.7) and PCR settings (see Table 4.8).

Table 4.7 PCR mix components to amplify hTFIIIC into pbiGBac vector

Component	Amount
5x High-Fidelity PCR Green buffer (ThermoFisher)	10 μL
dNTPs (10 mM)	1 μL
Cas forward primer (20 pmol/ μl)	1.25 μL
Cas reverse primer (20 pmol/ μl)	1.25 μL
DNA (10 ng/uL)	0.5 μL
Phusion polymerase	0.5 μL
Water	Up to 50 μL

Table 4.8 PCR settings to amplify hTFIIIC into pbiGBac vector. 35 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	30 s
Denaturation	98 °C	10 s
Annealing	69 °C	30 s
Extension	72 °C	15-30 s/kb
Final extension	72 °C	2x extension time

The enzyme was then deactivated by heating the mixture for 20 min at 80 °C. The next step involved the linearization of the pBIG1 vector; this was achieved by adding 1 μL of SwaI enzyme (NEB) and 1 μL of 10x buffer O (NEB) to a final reaction volume of 10 μL. The reaction was left to incubate overnight at 30 °C, followed by SwaI inactivation the next day for 20 min at 65 °C. To isolate linearized pBIG1 plasmids, the reaction was run on a 1% agarose gel, and the relevant band was excised and purified using the method previously outlined in the section on restriction-free cloning.

With the hTFIIIC gene inserts now amplified using CAS primers and the pBIG1 plasmid linearized, a Gibson assembly was performed. In this step, 33 ng of the linearized pBIG1 plasmid was mixed with a five-fold molar excess of each hTFIIIC gene. To 5 μL of this mixture, 15 μL of Gibson Assembly Master Mix was added and incubated at 50 °C for 60 min. The resultant mixture was used to transform Top10 chemically competent cells (ThermoFisher) according to the manufacturer's instructions. The transformed cells were then plated on agar containing Gentamicin. Selected colonies were cultured in 5 mL of LB medium supplemented with Gentamicin. The cultures were grown overnight, and plasmid DNA was subsequently isolated using a Miniprep Kit (ThermoFisher). The purified plasmids were digested with the SwaI restriction enzyme at 30 °C as described above. The sample was applied to a 1% agarose gel and subjected to electrophoresis for 45 min at 100 V. This step was essential for verifying the successful assembly of the genes. The plasmids that exhibited band patterns consistent with the anticipated simulations, as analyzed using SNAPgene software, were selected. These chosen plasmids were then sent to Eurofins Genomics for Sanger sequencing.

Table 4.9: Reaction mix to confirm successful gene assembly via SwaI test digestion

Component	Amount
plasmid (100 ng/µL)	1 µL
SwaI	0.5 µL
Buffer O (10x)	1 µL
Water	Up to 10 µL

To assemble the τA and τB genes located in the pBig1a and pBig1b vectors respectively, a linearized form of the pBig2 vector was required. For this, 1 µg of pBig2 plasmid was subjected to a restriction digestion reaction. The reaction setup included 1 µL of 10x CutSmart buffer (NEB), 1 µL of PmeI enzyme (NEB), and Nuclease-free water to a final volume of 10 µL. This mixture was incubated overnight at 37 °C. Post-digestion, the plasmid was purified using the same protocol as for the linearized pBig1.

36 ng of the pBig2 vector was combined with the insert DNA, which was used at a fivefold molar excess. The insert DNA included τA-containing pBig1a and τB-containing pBig1b. After initial mixture, the total volume was adjusted to 12.5 µl with Nuclease-free water, followed by the addition of 4 µL of 5x Isothermal Reaction Buffer (IRB) and 1 µL of PmeI enzyme (NEB). This mixture was incubated at 37 °C for 2 hours. Subsequently, 2.5 µL of a pre-prepared Gibson Assembly mix was added, which consisted of 8 µL of Taq DNA Ligase, 1 µL of T5 Exonuclease (diluted 1:30 in 1x IRB), and 1 µL of Phusion DNA Polymerase (10,000 units/ml). The reaction was then incubated for 60 min at 50 °C. The resulting product was used to transform Top10 chemically competent cells. Transformed cells were plated on agar containing chloramphenicol and incubated overnight at 37 °C. Selected colonies were then cultured in 3-5 mL of LB medium with chloramphenicol for 18 – 24 hours. Plasmid DNA was extracted from these cultures using a Miniprep kit (Thermo Fisher Scientific). As with pBig1, a test digestion was performed prior to sending the samples for sequencing (Eurofins Genomics).

4.5.2 Recombinant human τA subcomplex

Given that the human TFIIIC genes were already present in their respective pAceBac plasmids (as detailed in the 'Recombinant Human TFIIIC' section), Helga Grötsch performed the restriction-free cloning, detailed earlier, to insert a Twin-Strep-tag at the N-terminus of GTF3C3 (refer to Table 4.10 for details in the construct). Additionally, H.G successfully cloned GTF3C3 with a Twin-Strep-tag, GTF3C5 with a His-tag, and GTF3C6, each from their respective pAceBac vectors, into a pBig1 plasmid using the previously describe biGBac system method.

Table 4.10: Recombinant expression construct for τA subcomplex in insect cells

Gene	Tag
GTF3C3	Twin-strep-tag in N-terminus
GTF3C5	His-tag in N-terminus
GTF3C6	No Tag

4.5.3 Insertion of ybbR-tag in yeast TFIIICΔ593

Before starting my PhD project in the lab, former student Mathias Vorländer already worked on the pACEBac plasmids, which included all the yeast TFIIIC genes and their respective tags, as reported in reference [3]. Two constructs were aimed to be produced, the first construct having an insertion of the 11-amino acid ybbR tag between amino acids 286 and 287 of the τ95 subunit. The second construct involved placing the tag between amino acids 506 and 507 in the same subunit.

My initial approach involved performing an overlap extension PCR, wherein the primers were designed to include the ybbR tag sequence. For this, I used the τ95_Δ593 subunit construct in its pACEBac plasmid, where a stop codon had been introduced at position 593 (for more details see reference [37]). The specifics of the PCR mix preparation are detailed in Table 4.11 and the PCR set up in Table 4.12.

Table 4.11 PCR components used to incorporate an ybbR tag into the τ95_Δ593 subunit

Component	Amount
2x Mater Mix Q5 High-Fidelity (New England Biolabs)	12.5 µL
Forward primer (20 pmol/ µl)	1.25 µL
Reverse primer (20 pmol/ µl)	1.25 µL
Plasmid DNA (50 ng/uL)	1 µL
Water	Up to 25 µL

Table 4.12 PCR settings to insert an ybbR tag into the τ95_Δ593 subunit. 25 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	2 min
Denaturation	98 °C	20 s
Annealing	69 °C	20 s
Extension	72 °C	2 min 24 s
Final extension	72 °C	10 min

Due to the unsuccessful integration of the ybbR-tag via overlap extension PCR, Jonas Weidenhausen performed a Gibson assembly instead. For each construct, two DNA fragments were amplified: the first fragment from the start of the τ 95 gene until the ybbR insertion site (called insert) and the second fragment covering the pACEBac backbone and the remaining 3' end of the τ 95 gene including the ybbR insertion (called backbone). Each pair of fragments had 20 bp overlapping regions on both ends. The PCR reaction components and settings are outlined in Tables 4.13 and 4.14, respectively.

Table 4.13 PCR components used to incorporate an ybbR tag into the τ95_Δ593 subunit

Component	Amount
2x Mater Mix Q5 High-Fidelity (New England Biolabs)	12.5 µL
Forward primer (20 pmol/ µl)	2.5 µL
Reverse primer (20 pmol/ µl)	2.5 µL
Plasmid DNA (50 ng/uL)	1 µL
Water	Up to 25 µL

Table 4.14 PCR settings to insert an ybbR tag into the τ95_Δ593 subunit. 35 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	2 min
Denaturation	98 °C	20 s
Annealing	60 °C	20 s
Extension	72 °C	40 s (insert) 1 min 45 s (Backbone)
Final extension	72 °C	10 min

The resulting DNA fragments were purified from 0.8 % agarose gel using the Qiagen gel extraction kit following its manual. Subsequently, the fragments were assembled together using a standard Gibson assembly protocol. In this process, 50 ng of the backbone, was mixed with the insert DNA in a 1:3 molar ratio. The mixture was then brought up to 5 μL with water. This was followed by the addition of 15 μL of Gibson assembly master mix. The assembly reaction underwent a 30-min incubation at 50 °C. Afterward, 10 μL of this reaction mixture was used to transform DH5α competent cells (ThermoFisher). These cells were then cultured on LB agar plates supplemented with gentamicin and incubated overnight at 37 °C.

To screen for successful clones, I proceeded with a colony PCR protocol. Selected colonies were first diluted in 20 μL of dH2O. Using the same tip, the colonies were re-streaked onto fresh LB agar plates containing gentamicin to serve as a backup for the tested colonies. These plates were incubated overnight at 37 °C and then stored at 4 °C. The diluted colonies were used to prepare the PCR mix, following the specifications in Table 4.15, and the PCR was conducted according to the settings in Table 4.16.

Table 4.15 PCR components used to test positive clones with ybbR tag

Component	Amount
Green Phusion Mix (PEPCORE)	10 μL
Forward primer (0.5 μM)	1 μL
Reverse primer (0.5 μM)	1 μL
Colony dilution	1 μL
Water	Up to 20 μL

Table 4.16 PCR settings to test positive clones with ybbR tag. 25 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	5 min
Denaturation	98 °C	20 s
Annealing	69 °C	20 s
Extension	72 °C	45 s
Final extension	72 °C	3 min

To assess the successful incorporation of the ybbR tag into the two distinct constructs at positions 286-287 and 506-507, I analyzed each sample using gel electrophoresis. Specifically, 6 μL of the reaction product was loaded onto a 0.8% agarose gel, prepared with 1x TBE buffer and stained with 3 μL of SYBR Safe DNA stain (Life Technologies). Electrophoresis was carried out at 100 V for 30 min, allowing for the visualization of ybbR-tag integration. Following the confirmation of ybbR-tag presence in the colonies, the remaining 19 μL of the colony dilution (from previous step) was added to 10 mL of LB medium supplemented with gentamicin. This culture was then incubated at 37 °C with shaking to promote growth overnight. Subsequently, the target plasmid was isolated using a Miniprep Kit (ThermoFisher). The purified plasmid was sent to Eurofins Genomics for sequence verification.

To integrate the τ55 and τ131 genes, along with the new constructs τ95-286ybbR and τ95-506ybbR, into the pbiG1 plasmids, the inserts were amplified using primers specified in Table 4.17.

Table 4.17 Primers used for the amplification of τA genes into the pBiG1a vector.

pbiGBac vector	Gene	Primers
pbiG1a	τ55	CasI forward and CasI reverse
	τ95-286ybbR or τ95-506ybbR	CasII forward and CasII reverse
	T131	CasIII forward and CasV reverse

After the successful insertion of the τA genes into pBiG1a, the τB genes already inserted in pBiG1b were then assembled into pbiG2. This process followed the protocol outlined in Section 4.3.1, 'Recombinant human TFIIIC.

4.6 Protein expression and purification

4.6.1 Recombinant human τA subcomplex
To produce the human τA baculovirus, the plasmid outlined in Section 4.5.2, 'Recombinant Human τA Subcomplex' was used. The production of this baculovirus followed established protocols, which were adapted for our specific laboratory needs by H.G. and M.V. Bacmid DNA preparation was executed by H.G., followed by transfection into insect cells. On the first day Agar plates containing 1 mM IPTG and a selection of antibiotics (Kanamycin – 50 µg/mL, Ampicillin - 100 µg/mL, Tetracycline - 10 µg/mL, Gentamicin - 7 µg/ml) were coated with 200 µL of X-Gal (200 µg/mL) and allowed to dry for subsequent bacterial transformation. EmBac YFP cells underwent electroporation with 1-2 µL of the plasmid construct. These cells were then cultured in SOC medium at 37 °C under agitation for recovery. The following day, cells were plated on the agar plates prepared previously (day one) and incubated at 37 °C. By the third day, these plates were evaluated to identify and select white colonies. Selected colonies were then expanded in 3 mL of SOC medium, which was supplemented with the same antibiotics as in the agar plates.

To isolate the bacmid DNA, the cell medium was first centrifuged for 5 min at 4000 g, pelleting the cells. The pellet was then resuspended in 250 µL of Buffer P1 (Qiagen miniprep kit). This was followed by the addition of 250 µL of Buffer P2 and 350 µL of Buffer N3, after which the sample was centrifuged for 10 min at 17,900 g. The supernatant was then treated with 700 µL of isopropanol and centrifuged again for 10 min to precipitate the DNA. The DNA pellet was washed with 200 µL of 70% ethanol and centrifuged for 5 min. After air-drying the pellet under a hood, it was resuspended in 25 µL of water for subsequent use.

For V_0 virus production, 3 mL of SF21 cells at 0.3×10^6 cells/mL were aliquoted into each well of a 6-well plate, including one additional well for a cell-only control. After allowing the cells to adhere for 15 min, the bacmid DNA (15 µL DNA at 900 ng/µl) was mixed with 100 µL of warm Sf900 III media (mix 1). Separately, a transfection solution was made by mixing 5 µL of the X-tremeGENE^{TM} HP DNA transfection Reagent High performance (Sigma-Aldrich) with 50 µL of media per sample. Then, 55 µL of this reagent mix was combined with the bacmid DNA (mix 1) and incubated at room temperature for 15 min. The resulting mixture was added dropwise to each well. The 6-well plate was then incubated at 27 °C for 48-72 h, wrapped in aluminum foil without shaking. Harvesting was typically done on the morning of the third day, approximately 66 h post-transfection. Post-harvest, the virus production was checked under the microscope for reduced cell density, enlarged cells, and visible cell debris. Additionally, green fluorescence was also evaluated. Finally, the supernatant was collected and stored in a 15 mL falcon tube in the fridge for subsequent use.

I continue with the virus production and subsequent steps. For production of the V_1 virus, 0.5 to 3 mL of the V_0 virus supernatant was used. This was used to infect 25 mL of SF21 cell culture at a density of 0.5

million cells/mL. Over the subsequent three days, cell parameters were monitored. An optimal infection is marked by the cells undergoing a single doubling; failure to achieve even one division may indicate an excessively high viral titer, elevating the mutation risk. Post-infection, cell division should stop. Daily observations of Green Fluorescence Protein (GFP) fluorescence were made when cellular division halted, which is indicative of baculovirus expression. Typically, the virus titer peaks 72 hours after the cessation of growth. The culture was then centrifuged at 600 g for 10 min, and the supernatant, designated as V_1, was transferred to a 50 mL Falcon tube and stored in the dark. The cell pellet was washed once with PBS and subsequently stored at -80°C for use in expression analysis.

Prior to scaling up τA expression, a virus titration was necessary. Both SF21 and High Five cells were assessed for expression levels. For this, 25 mL cell cultures at a density of 0.5×10^6 cells/mL were infected with virus dilutions of 1:500, 1:1000, and 1:3000. Cell viability and growth were monitored in the following days. Ideally, cells should divide once before growth arrest. Two days after stop division, the cells, which typically have a cell viability of 95-100%, were harvested. They were centrifuged at 600 g for 5 min, wash with 1x PBS and then frozen. The small pellet was then subjected to a pull-down assay to verify protein expression. If the protein expression was satisfactory, a larger culture was initiated, employing the most effective virus dilution from the titration.

The τA subcomplex was expressed in High Five cells, using a virus titration of 1:1000. One liter of cell culture at a density of 1×10^6 cells/mL was infected and incubated until the cell viability reached 90-95%, typically within 72 h. Then, cells were harvested by centrifugation at 500 g for 10 min at RT. The cell pellets were subsequently washed with 1x PBS and re-centrifuged at the same speed and duration. The final pellets were quickly frozen in liquid nitrogen and preserved at -80 °C for later use.

Cell pellets from a 6-L culture were resuspended in lysis buffer (20 mM HEPES pH 7.5, 4 mM β-mercaptoethanol, 500 mM NaCl, 2 mM MgCl2, 10% glycerol), using 3 mL of buffer per gram of pellet. Protease inhibitor cocktail tablets without EDTA (Sigma-Aldrich) were added to the suspension at a ratio of 1 tablet per 20 g of cells, along with 4 μL of PEPCORE Benzonase per 50 mL of buffer and 500 μL of DNase I in total (10 mg/mL). This mixture was stirred contiously for 1-2 hours in a metal beaker placed on ice inside a cold room for homogenization. Thawed cells were sonicated for 3 min at 40% amplitude with intervals of 10 seconds on and 10 seconds off, while being kept in an ice-water bath to prevent overheating. Post-sonication, the lysate was ultracentrifuged at 45,000 rpm for 1 h at 4 °C. The clarified supernatant was then filtered using 1.2 μm syringe filters. This filtrate was applied to a 5 mL Strep-Tactin® Superflow® high-capacity column (IBA), pre-equilibrated with strep-wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT, 5% glycerol), at a flow rate of 0.5 mL/min. After loading the sample, the column was washed with 35 mL of the same strep-wash buffer and then eluted with 30 mL of strep-elute buffer containing 50 mM biotin. The eluate was subsequently loaded onto a pre-equilibrated Capto HiRes Q 5/50 column using a super loop, washed with 10 mL of Capto-A buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT), and eluted with a linear gradient of 0 to 70 % followed by a step gradient to 100 % of Capto-B buffer (20 mM HEPES pH 7.5, 1 M NaCl, 2 mM DTT). Fractions displaying a conductivity between 38.1 mS/cm and 48.07 mS/cm were pooled, subjected to buffer exchange in Capto-A buffer, concentrated to a final concentration of 9.65 mg/mL, and then flash-frozen for storage.

4.6.2 Recombinant human TFIIIC

I carried out the baculovirus production of hTFIIIC in accordance with the protocol described for the recombinant τA subcomplex (Section 4.6.1), substituting SF21 cells for High Five cells. Moreover, I personally performed the purification of hTFIIIC.

Frozen SF21 cell pellets were resuspended in 3x lysis buffer comprising 20 mM HEPES pH 7.5, 500 mM NaCl, 2 mM MgCl2, 0.1% NP-40, 10% glycerol and 0.25 mM DTT. For every 20 g of cell pellet, one SigmaFast EDTA-free protease inhibitor tablet, 4 μL of Benzonase per 50 mL of buffer, and a 1:1000 dilution of DNase I were added. The cells were stirred gently for 2 h at 4°C. Subsequently, cell disruption was achieved through a 3 min sonication step, with 10 s intervals, at 40 % amplitude. The lysate was then clarified by ultracentrifugation at 30,000 g for 1 h at 4 °C. The supernatant obtained was incubated with Anti-FLAG M2 agarose beads (Sigma-Aldrich) for 2 h on a roller at 4 °C. The bead-bound complex was loaded onto an Econo-Column Chromatography column and subjected to an extensive washing: first with 30 column volumes (CV) of the lysis buffer, followed by 25 CV of a wash buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10 % glycerol, 0.1% NP-40, and 0.25 mM DTT). Elution was performed with 2 CV of a specific elution buffer (wash buffer with FLAG peptide at 0.2 mg/mL) for 30 min at 4 °C. The eluted fractions were then adjusted with wash buffer 2 (20 mM HEPES pH 7.5, 100 mM NaCl, 10 % glycerol, 0.1 % NP-40, 0.25 mM DTT) to bring the NaCl concentration down to 200 mM. This adjusted sample was loaded onto a pre-equilibrated Capto HiRes Q 5/50 column (Cytiva) in buffer A (20 mM HEPES pH 7.5, 200 mM NaCl, 5 % glycerol, 5 mM DTT). A linear gradient of 0 to 100 % buffer B (20 mM HEPES pH 7.5, 1 M NaCl, 5 % glycerol, 5 mM DTT) was applied over 70 ml for elution. SDS-PAGE was used for analyzing peak fractions. TFIIIC-containing fractions (normally between 25.5 - 30.5 mS/cm) were pooled, concentrated to 2 mg/ml, buffer-exchanged into hTFIIIC storage buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 5 mM DTT), flash-frozen, and stored at −80 °C or directly used for cryo-EM sample preparation. The integrity of hTFIIIC was assessed using mass photometry, as detailed in section 4.4.

4.6.3 Recombinant yeast TFIIICΔ593 and yeast ybbR-TFIIIC-Δ593

In our laboratory, two pbiB2 plasmids containing two variants of the yeast TFIIIC were used for baculovirus production. The initial variant comprised the entire yeast TFIIIC genes, substituting τ95 subunit with τ95-Δ593 mutant. In this configuration, a stop codon replaced the amino acid at position 594, resulting in the construct named TFIIIC-Δ593, developed by M.V and A.J. The second variant is a construct with ybbR-TFIIICΔ593 as detailed in section 4.5.3. I followed the experimental procedures described in section 4.6.1.

Expression of yTFIIICΔ593 was carried out in High Five cells, while ybbR-yTFIIICΔ593 was expressed in SF21 cells. The purification protocol applied to these two complexes was similar to that used for the human τA subcomplex, including purification steps and buffer compositions, as detailed in section 4.6.1. The main difference in the purification was the replacement of the 5 mL Strep-Tactin® Superflow® high-capacity column with 2 mL (50 % slurry) of Strep-Tactin Sepharose™ High Performance beads. After anionic exchange chromatography, fractions containing TFIIIC, particularly within the conductivity range of 25 to 32.32 mS/cm, were collected and stored at −80 °C until use.

4.7. Single Molecule Fluorescence Microscopy applied to yTFIIIC

4.7.1 ybbR-TFIIICΔ593 labeling with cy5 and cy5.5 dyes

To label the ybbR-yTFIIICΔ593 complex, which purification is detailed in section 4.6.3, I used reagents generously provided by Simone Höfler. This included the Sfp enzyme at a stock concentration of 1.4 mg/ml (49.6 μM) and the fluorescent CoA derivatives, CoA-Cy5 at 0.4 mM and CoA-Cy5.5 at 1.75 mM.

This protocol can be divided into two different steps: labeling optimization and preparative labeling. During the optimization phase, a variety of conditions outlined in Table 4.18 were tested at 25 °C using a labeling buffer (50 mM HEPES-KOH pH 7.5, 100 mM NaCl, 20 mM MgCl2 and 5 mM DTT). Condition 7 was also assessed at 37 °C to compare the effects of temperature on labeling efficiency (reaction 10). Reactions were set up in 20 μL volumes, with half dedicated to SDS-PAGE analysis and the other half to functionality assessment. The SDS-PAGE gels were initially scanned on a Typhoon instrument using the Cy5 channel before Coomassie staining. Complex integrity was verified using EMSA as per the protocol in section 4.2.2. Cy5 and Cy5.5 signal integration for quantification was performed using image analysis software, such as Fiji/Image J.

Table 4.18 Experimental conditions evaluated for optimal labeling of ybbR-yTFIIICΔ593

Component	Reaction 1	Reaction 2	Reaction 3	Reaction 4	Reaction 5	Reaction 6	Reaction 7	Reaction 8	Reaction 9
ybbR-yTFIIICΔ593	5 μM	5 μM	5 uμM	5 μM	5 μM	5 μM	5 μM	5 μM	5 μM
Sfp	0.1 μM	0.1 μM	0.1 μM	1 μM	1 μM	1 μM	2.5 uM	2.5 μM	2.5 μM
CoA-Cy5 or CoA-Cy5.5	10 μM	10 μM	10 μM	10 μM	10 μM	10 μM	10 μM	10 μM	10 μM
Incubation duration	30 min	60 min	90 min	30 min	60 min	90 min	30 min	60 min	90 min

Upon identifying an optimal labeling condition yielding a functional and effectively labeled yTFIIIC, I continue with the preparative labeling step. This involved scaling the selected condition to a 500 μL reaction volume. After the reaction, excess dye was eliminated via buffer exchange using a PD-10 desalting column (Sigma-Aldrich). The eluates, collected in approximately 100 μl fractions, were analyzed by SDS-PAGE to detect any residual dye. Fractions with minimal dye contamination were combined and further purified with a fresh PD-10 column, followed by another SDS-PAGE for analysis.

The absorbance of the pooled fractions was measured at 280 nm and 646 nm with a Nano-Drop spectrophotometer before aliquoting and storage at -80 °C. This measurement was crucial for assessing both the labeling efficiency and the concentration of the protein. Furthermore, an additional SDS-PAGE step was conducted to accurately determine and adjust the concentration of the labeled protein. Three different quantities of unlabeled protein and equivalent amounts of Cy5- and Cy5.5-yTFIIICΔ593

complexes were loaded onto an SDS-PAGE gel. Band intensities of the unlabeled protein were quantified using Image Lab software (Bio-Rad), and a linear regression model was applied to these results.

4.7.2 DNA template generation

To prepare the His_tH(GUG)E2 gene for single-molecule fluorescence microscopy (SMFM), adaptor sequences were introduced at both termini. The 5' overhang was designed to be complementary to a sequence tagged with Cy3 dye at its 5' termini. The 3' overhang was designed to align with a sequence facilitating the immobilization of the tRNA gene sequence. Nusrat Shahin Qureshi designed the primers required for introducing these specific overhangs. Subsequently, I performed an analytical PCR gradient to determine the optimal annealing temperature and then proceed with a preparative PCR to maximize the yield of the intended template.

The analytical PCR gradient was carried out using a master mix, prepared for 10 reactions corresponding to each temperature tested, following the parameters and settings as outlined in Tables 4.19 and 4.20.

Table 4.19 PCR components used for PCR gradient to modify His_tH(GUG)E2 for SMFM

Component	Amount for 10 reactions
2x Green Phusion Mix (PEPCORE)	125 µL
Forward primer (0.5 µM)	12.5 µL
Reverse primer (0.5 µM)	12.5 µL
DNA template (10 ng/µl)	5 µL
Water	Up to 250 µL

Table 4.20 PCR settings for PCR gradient to modify His_tH(GUG)E2 for SMFM. 35 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	1 min
Denaturation	98 °C	10 s
Annealing	Gradient*	30 s
Extension	72 °C	20 s
Final extension	72 °C	2 min

*Gradient applied: 53.4, 53.8, 55, 57.1, 59.7, 62.3, 64.7, 67.3, 70, 72 °C

The samples were loaded onto a 1% agarose gel, prepared with 1X TBE and 3 µl of Sybr Green, and then run for 30 min at 100 V. Band intensities on the gel were visually compared to identify the optimal annealing temperature.

The preparative PCR, aimed at producing a sufficient template volume equivalent to 20 reactions, was performed in accordance with the parameters and settings detailed in Tables 4.21 and 4.22.

Table 4.21 PCR components used for preparative PCR to modify His_tH(GUG)E2 for SMFM

Component	Amount for 20 reactions
2x Green Phusion Mix (PEPCORE)	250 µL
Forward primer (0.5 µM)	25 µL
Reverse primer (0.5 µM)	25 µL
DNA template (10 ng/µl)	10 µL
Water	Up to 500 µL

Table 4.22 PCR settings use for preparative PCR to modify His_tH(GUG)E2 for SMFM. 35 rounds of amplification were used

Stage	Temperature	Duration
Initial denaturation	98 °C	1 min
Denaturation	98 °C	10 s
Annealing	59.7 °C	30 s
Extension	72 °C	20 s
Final extension	72 °C	2 min

The samples were run on a 1% agarose gel, prepared with 1X TBE and 3 µL of Sybr Green, for 30 min at 100 V. Subsequently, the DNA bands were extracted from the gel using the Gel Extraction Kit (Qiagen), following the manufacturer's instructions.

The purified fragments were buffer exchanged using SMFM DNA buffer (10 mM Tris-HCl pH 7.5, 20 mM KCl) using an Amicon Ultra-0.5 Centrifugal Filter 3 kDa MWCO Millipore. This step was critical for ensuring the integrity of subsequent binding experiments in single-molecule fluorescence microscopy (SMFM). DNA was buffer exchange 3 times using the same centrifugal filter. The buffer exchange was performed three times using the same filter unit. To assess its effectiveness, the flow-through was analyzed at 280 nm using a Nano-Drop spectrophotometer. During the final buffer exchange, the sample volume was carefully adjusted to reach a target range of 30–50 µL at around 1 µM.

4.7.3 Total internal reflection fluorescence (TIRF) analysis of ybbR-yTFIIICΔ593 and DNA

The TRIF experiments were performed by Anastasiia Chaban. For TIRF microscopy preparation, the DNA template (modified His_tH(GUG)E2 gene) detailed in Section 4.7.1, was incubated with two distinct oligonucleotides. A biotinylated oligonucleotide was used to anneal with the 3' overhang, facilitating the immobilization of the tRNA gene. Simultaneously, a Cy3-tagged oligonucleotide was added to hybridize with the gene's 5' overhang. To accomplish this, all the components were combined in the SMFM DNA buffer (composition detailed in the previous section) and incubated for 5 min at 65 °C, followed by a gradual cooling step to room temperature over a period of 1 h. The specific components and their respective concentrations are listed in Table 4.23.

Table 4.23: Components added to the modified tRNA gene for TIRF analysis

Component	Amount	Final concentration
tRNA gene (1.8 µM)	1.09 µL	0.1 µM
Cy3-oligonucleotide (1 µM)	2 µL	0.1 µM
Biotinylated oligonucleotide (0.25 µM)	8 µL	0.1 µM
SMFM DNA buffer	Up to 20 µL	-

The reaction mixture was sequentially diluted from 100 nM to 50 pM through two dilutions steps using SMFM reaction buffer, which consists of 20 mM HEPES pH 8, 100 mM KCl, and 2 mM MgCl$_2$. The final dilution step yielded a total volume of 200 µl, which was then maintained on ice for subsequent use.

To continue with the TIRF experiment, neutravidin at a final concentration of 100 µg/mL diluted in a buffer containing 10 mM Tris-HCl pH 8.0, 50 mM NaCl, was first inyected into the in-house microfluidic channel and incubated for 5 min to promote surface binding. Following this, the channel was washed with SMFM reaction buffer to remove any non-adherent neutravidin. Subsequently, 200 µl of the 50 pM DNA solution was applied to the channel and left to incubate for 10 min, allowing the DNA to bind to the neutravidin-coated surface. After DNA application, the channel was further cleansed with Oxygen Scavenger (OSC) wash buffer (2.5 mM PCA, 2 mM TSY, 50nM PCD, 0.25% Biolipidure 203, 0.25% Biolipidure 206) to eliminate any unbound DNA. Finally, once the sample was placed on the microscope and the recording initiated, an OSC chase solution (2.5 mM PCA, 2 mM TSY, 50nM PCD, 0.25%

Biolipidure 203, 0.25% Biolipidure 206) containing 100 nM ybbR-TFIIICΔ593-Cy5 was injected 10 seconds into the same microfluidic channel to track the interaction dynamics. The TIRF experiment was conducted over 10 min. Subsequently, the acquired traces were processed using SPARTAN software package ran with MATLAB, with a focus on visually identifying and selecting traces that displayed clear complete photobleaching, thereby ensuring the exclusion of instances involving more than one DNA molecule.

4.8 Cryo-EM studies of Human τA subcomplex

4.8.1 Stabilization of Human τA subcomplex for cryo-EM

I carried out a detergent and grid screening on the human τA sample purified as described in Section 4.6.1, using the Talos™ Arctica™ microscope (ThermoFisher). The frozen sample was thawed on ice and diluted from its original concentration of 9.65 mg/mL(53.9 μM) to 1.2 mg/mL (6.6 μM), 0.6 mg/mL (3.3 μM) and 0.48 mg/ml (2.7 μM). This dilution process was executed using Capto-A buffer, the composition of which is detailed in Section 4.6.1. For grid preparation, a plasma cleaner with a 10% argon and 90 % oxygen mixture was employed for 2 min and 30 s. The vitrobot Mark IV (ThermoFisher Scientific) was set to 6 °C with 100 % humidity for this process. A range of detergents were applied immediately prior to plunge freezing the samples. The specifics of the grids and detergents used, including their final concentrations, are comprehensively listed in Table 4.24.

Table 4.24: Selection of grids and detergents for stabilizing human τA in cryo-EM studies

Type of grid	Detergents
Quantifoil R2/1 Cu 200	0.5 % Octyl-beta-Glucoside (OG)
Quantifoil R2/2 Cu 400	1 % Lauryl Maltose Neopentyl Glycol (LMNG)
Quantifoil R1/4 Au 200	-
Quantifoil R1/2 Au 200	-

4.8.2 Human τA cryo-EM data collection and processing

To further assess the integrity of the human τA sample, a small dataset was acquired using the Talos™ Arctica™ microscope from ThermoFisher, operating at 200 keV and equipped with a Falcon3 detector. A total of 1344 image stacks, each comprising 10 frames, were collected in linear mode using SerialEM [125]. This setup included a magnification of 120,000 x, translating to a physical px size of 1.237 Å.

The raw movies were initially processed on the fly using the WARP software [126]. Particle picking was carried out with the same program, using the BoxNet2_20180918 model without any retraining. Following this, image stacks underwent preprocessing in RELION 3.1.3 [127], utilizing its integrated MotionCor2 [128] feature for motion correction. The Contrast transfer function (CTF) parameters were determined using Gctf [129]. From the WARP-selected particles, a total of 188,673 were extracted, each with a box size of 384 px. These particles were then imported into cryoSPARC [130] for all subsequent data processing steps.

Following the confirmation of the integrity of the human τA subcomplex from the data acquired using the Talos microscope, a new dataset was collected on the Titan Krios G3 (ThermoFischer Scientific)

operated at 300 keV. This microscope was equipped with an energy filter and a Gatan K3 detector. A total of 8779 image stacks were recorded, each comprising 40 frames with an exposure rate of 1.34e/Å^2/frame. A magnification of 130,000 x, corresponding to a px size of 0.645 Å, within a defocus range of -0.9 to -1.9 was used. The pre-processing steps used for the Talos data were also applied to this dataset. Additionally, to obtain a greater Number of particles compared to those selected by WARP, a subset of particles derived from these initial ones was utilized to train neural networks for particle picking in TOPAZ [131]. These new particles were then imported into cryoSPARC for the remaining data processing steps.

4.9 Structure determination of human TFIIIC and TFIIIC-DNA complex

4.9.1 Sample preparation of hTFIIIC and hTFIIIC-DNA complex
To determine the cryo-EM structure of hTFIIIC both without DNA and in complex with DNA, I used freshly purified hTFIIIC. Plunge-frozen samples were utilized in biochemical and biophysical experiments only.

To prepare "hTFIIIC without DNA" sample for cryo-EM, the concentration of hTFIIIC was adjusted from 3.17 µM (2 mg/ml) to 1.59 µM. The buffer composition was similar to that of the purification buffer exchange (see Section 4.6.2), with modifications in NaCl concentrations at 150, and 200 mM to explore different conditions. Samples with NaCl concentrations below 200 mM underwent dialysis using 10 kDa Slide-A-lyzer Mini units (Sigma-Aldrich) at 4 °C for 1 h. Before plunge-freezing, OG detergent was added to the sample to prevent air-water interface interactions and ensure uniform distribution across the Ultrafoil R2/2 Au 200 grid, similar to the protocol applied for the human τA subcomplex sample preparation. Detergent concentrations were varied, tested at final concentrations of either 0.5 % or 1 %. The plasma cleaner and Vitrobot settings were kept consistent with those used for the human τA subcomplex.

Prior to the hTFIIIC-DNA complex sample preparation, EMSA and MP assays (refer to sections 4.2 and 4.4, respectively) were performed with the human TRR-TCT3-2 gene to verify the activity of recombinant hTFIIIC. For complex formation, 1.59 µM hTFIIIC was incubated with an equal molar concentration of tRNA DNA oligonucleotide at room temperature for 10 min. The hTFIIIC-DNA mixture was then subjected to a Zeba Spin desalting column (ThermoFisher Scientific), pre-equilibrated with a 20 mM HEPES pH 8.0, 5 mM MgCl$_2$, and 5 mM DTT buffer, while varying KCl concentrations (125, 175, 225, and 250 mM) to evaluate the effect on complex stability. For electron microscopy, the grid preparation—including grid type, detergent concentration, plasma cleaning, and Vitrobot settings for vitrification—was consistent with the protocols used for the hTFIIIC samples without DNA, ensuring uniformity in our comparative analysis.

4.9.2 Electron microscopy and data processing of hTFIIIC and hTFIIIC-DNA complex
To avoid redundancy, this section will focus exclusively on the data processing of the hTFIIIC-DNA complex. For data processing and collection details of the "hTFIIIC without DNA" sample, see Figure 2.9 and Table A1-A2, respectively.

Cryo-EM data for the hTFIIIC-DNA sample was acquired using a Titan Krios G3 microscope (Thermo Fisher Scientific) at 300 keV, equipped with a Gatan K3 detector and energy filter. The collection was set to a magnification of 105,000x (physical pixel size of 0.822 Å). I collected 11,025 image stacks, each comprising 40 frames, in counting mode. The stacks were exposed to a total electron dose of 42.8 e/Å^2, across a defocus range of 0.7 to 1.7 μm, using SerialEM for automated data acquisition.

Initial preprocessing of micrographs was performed in WARP, utilizing the BoxNet2_20180918 model for particle picking. Following this, the micrographs were processed in RELION 3.1.3, applying MotionCor2 for motion correction and Gctf for CTF estimation. From this preprocessing, a set of 530,171 particles, extracted with a box size of 480 px, was imported to cryoSPARC 3.3.2 for subsequent classification and refinement steps. In cryoSPARC, 2D classification delineated two particle sets corresponding to τA and τB-DNA subcomplexes. Particles that contributed to high-resolution 2D classes for these two sets were used to train a neural network (conv127 model) in TOPAZ to improve particle picking. For τA, first and second round of TOPAZ-based particle picking yielded 280,410 and 365,063 particles, respectively. The last picked particles underwent a series of three heterogeneous refinement steps, the best class containing 55,079 particles was used for a non-uniform refinement, achieving a resolution of 3.5 Å. A similar approach was adopted for the τB-DNA subcomplex. The initial TOPAZ picking resulted in 517,116 particles, with a second round this number increased to 762,590. These particles were then classified and refined, culminating in a subset of 99,217 particles with a resolution of 3.3 Å using a non-uniform refinement job. Further particle classification was performed using cryoDRGN [132]. Particle sets were downsampled to 256 × 256 px. A neural network with three hidden layers, each layer having 512 neurons per layer for encoder and decoder, and an 8-dimensional latent space, trained over 50 epochs. The trained network's cluster averages were used to reconstruct cryo-EM density maps. Based on these maps, 35,379 particles were selected for non-uniform refinement in cryoSPARC. Additionally, local refinement of each monomer in the τB-DNA map was performed, achieving a resolution of 3.2 Å.

4.9.3 Model building, refinement, analysis and validation

I used the software Coot [133] for model building, with AlphaFold [134] predicted structures of hTFIIIC subunits (TFIIIC220: AF-Q12789-F1, TFIIIC110: AF-Q8WUA4-F1, TFIIIC102: AF-Q9Y5Q9-F1, TFIIIC90: AF-Q9UKN8-F1, TFIIIC63: AF-Q9Y5Q8-F1, TFIIIC35: AF-Q969F1-F1) obtained from the AlphaFold Database. These structures were placed into the density maps using ChimeraX [135]. A B-DNA model was positioned within the density, applying self-restraints. AlphaFold-multimer [136] was employed for regions of low resolution. Model refinement was iteratively done in PHENIX [137], coupled with manual adjustments in Coot. The refined model underwent validation with MolProbity [138] and surface interaction areas were quantified using PISA [139]. Protein-DNA interactions were examined using NUCPLOT [140]. For τA, model building was based on cryo-EM density from the TFIIIC-DNA reconstituted sample. τA particles from the DNA-unbound sample, refined to a 3.8 Å resolution, exhibited a preferred orientation at the air-water interface. Rigid-body fitting of the τA model from the "hTFIIIC-DNA sample" into the τA density from "hTFIIIC without DNA sample" revealed no significant conformational changes. Structural homologs of TFIIIC220 domains were identified using DALI [141] and the DNA geometry was analyzed with Curves+ [142].

4.9.4 τA to τB-DNA particle distance at a single molecule level

For the particle distance pair analysis, Luis Hauptmann utilized .star files from RELION, containing data from the final Non-unifom refinement of τA and τB-dimer subcomplexes derived from the 'hTFIIIC with DNA' sample. The τB-dimer and τA particle names and coordinates were extracted from each micrograph and arranged into a distance matrix (size NτB × NτA). Particle pairing was executed by identifying the minimum distance in the matrix, saving the corresponding particle pair, and then removing these particles to prevent re-pairing. This process was iterated until all particles were paired. This step was repeated for each micrograph, generating a new distance matrix each time. The distances for all pairings across micrographs were compiled and plotted in a histogram. To differentiate between linked and random particle pairs, a control set with randomized particle coordinates was created using the same particle sets from the micrographs. The distances for these randomized pairs were calculated in the same manner and compared to the actual measured distances of the particles.

4.10 Structure determination of yeast TFIIIC-DNA complex

4.10.1 Sample preparation of DNA-yTFIIICΔ593 complex

To reconstitute the TFIIICΔ593- DNA complex, a similar approach used to investigate the interaction of hTFIIIC with its tRNA gene was applied (see section 4.9.1). In the mass photometry assay, buffers with 20 mM HEPES pH 8.0, 2 mM $MgCl_2$, 5 mM DTT, and varying KCl concentrations (150, 175, 200, and 225 mM) were tested. More details about the MP experiment can be found in section 4.4. All the buffers tested in the MP experiment with TFIIICΔ593 and His_tH(GUG)E2 gene were further used for cryo-EM grid preparation. To form the complex, 1.97 µM of yTFIIICΔ593 was combined with an equimolar amount of tRNA DNA oligonucleotide. Subsequently, the mixture was processed through a Zeba Spin desalting column (ThermoFisher Scientific). The parameters for grid type, detergent concentration, plasma cleaning, and Vitrobot settings for vitrification were maintained as per the hTFIIIC experiment protocols, except where specifically noted otherwise.

Additionally, to stabilize the DNA-yTFIIICΔ593 complex with transcription factors such as yTFIIIB and fpt1, a series of buffers were tested. The experiments maintained a constant buffer composition of 20 mM HEPES pH 8.0, 2 mM MgCl2, and 5 mM DTT, while altering KCl concentrations ranging from 75 to 200 mM. For an in-depth comparison of these specific conditions across distinct sample preparations used in collecting five different datasets, refer to Table 4.25.

Table 4.25: Comparison of sample preparation components across the five distinct datasets

Datasets	Salt concentration	*DNA oligonucleotide	Additional transcription factor	Detergent concentration
Dataset1	200 mM KCl	85 bp	No	0.1 % OG
Dataset2	175 mM KCl	120 bp	Brf1-TBP fusion protein, BDP1 and fpt1	0.05 % OG
Dataset3	150 mM KCl	120 bp	fpt1	0.05 % OG
Dataset4	75 mM KCl	120 bp	Brf1-TBP fusion protein, BDP1 and fpt1	0.05 % OG
Dataset5	75 mM KCl	120 bp	Brf1-TBP fusion protein	0.05 % OG

*Two versions of the His_tH(GUG)E2 gene fragment, with a length of 85 bp (-9 to +86) and 120 bp (-43 to +86), were used. The 120 bp variant is distinguished by an additional 35 bp segment containing a TATA box mutation, enabling a strong yTFIIIB binding in the 5' upstream region.

4.10.2 Electron microscopy and data processing

To obtain high-resolution structures of the yTFIIICΔ593-DNA complex, an initial dataset (Dataset 1) comprising 19,047 image stacks, each with 40 frames, was acquired (refer to Table 4.25 for detailed composition of the sample). Data collection was performed on a Titan Krios G3 electron microscope (ThermoFischer Scientific), operating at an acceleration voltage of 300 keV. This microscope is equiped with an energy filter and a Gatan K3 direct electron detector. Imaging was conducted under a total electron dose of 39.6 electrons per $Å^2$ and a defocus range set from 0.7 to 1.7 µm. The chosen magnification of ×105,000 resulted in an effective px size of 0.822 Å.

The processing of the dataset, starting with 1,004,005 particles picked by WARP, was divided into three main classes based on the initial analysis: τB-DNA dimer, τB monomer, and τA-DNA subcomplexes. From each class, 2D classifications were used to create initial ab initio maps. Together with two 'junk' classes—generated by a few random selected particles—a step of heterogeneous refinement was performed in cryoSPARC to further sort the entire set of WARP-picked particles. For a comprehensive depiction of the data processing steps, see Figure 3.4.

During the processing of the τB-DNA dimer class, following the initial heterogeneous refinement step, a total of 199,781 particles were classified as dimer particles. These particles were then subjected to two additional rounds of heterogeneous refinement using the previously 'junk' classes. Then, a round of 2D classification was performed to select particles suitable for the initial TOPAZ training and picking within RELION. This approach yielded a new dataset of 336,587 particles. Further sorting of these particles was achieved through two rounds of heterogeneous refinement. After these two steps a map, containing 122,607 particles, was refined using non-uniform refinement in cryoSPARC. However, this map showed a preferred orientation, evident in Figure 3.4 (bottom left). Due to this preferred orientation, a second round of TOPAZ training and picking was not conducted. The focus was then shifted to the analysis and processing of the other two classes.

In the processing of the τB-DNA monomer class, the initial heterogeneous refinement step resulted in 293,371 particles classified as monomer particles. These particles underwent two further rounds of heterogeneous refinement using the previously designated 'junk' classes. Subsequently, a 2D classification step was carried out to select particles for the first round of TOPAZ training and picking in RELION, leading to a newly set of 589,216 particles. These particles were then processed through two additional rounds of heterogeneous refinement. Following this, the map representing the τB-DNA monomer underwent non-uniform refinement. The particles contributing to this map were then re-imported into RELION for a second round of TOPAZ training and picking. This process culminated in a final dataset comprising 546,857 particles, which were classified iteratively through two rounds of heterogeneous refinement. This led to a τB-DNA monomer map containing 258,272 particles with a resolution of 3.21 Å. For additional details and visual representation of this process, refer to Figure 3.4, specifically the bottom middle section.

For the τA-DNA subcomplex, the last distinctive class in this dataset, the processing began differently. The initial 2D classes, which were utilized to create the ab initio map for the first global heterogeneous refinement step, were directly applied for the first TOPAZ training and picking process. This method was chosen because using particles after heterogeneous refinement provided fewer particles for training. The first TOPAZ training and picking step generated a new set of 271,890 particles. A specific 2D classification step was carried out, leading to the selection of 26,803 particles. These selected particles were then used for the second TOPAZ training and picking step, resulting in a substantial increase in the particle count to 750,337 particles. From these particles, 50,000 random particles were used to create new 'junk classes'. Additionally, to improve the initial τA-DNA map before proceeding with heterogeneous refinement using the newly selected particles, a round of 2D classification was performed. This step focused on identifying and selecting classes that closely matched the τA-DNA class, which were then used for subsequent ab initio reconstruction and non-uniform refinement. Utilizing these updated maps for the heterogeneous refinement step, the τA-DNA map containing 125,533 particles achieved a resolution of 6.54 Å after non-uniform refinement. Additional details and a visual representation of this process can be found in Figure 3.4, especially in the bottom right section.

To improve the resolution of the τB-DNA and τA-DNA subcomplexes, I collected four more datasets across different sessions. The Titan Krios G3 microscope (ThermoFischer Scientific), operating at 300 keV, was utilized for data collection. This microscope is equipped with a Gatan K3 detector and an energy filter. The magnification setting was at ×105,000, yielding a physical pixel size of 0.822 Å. Image stacks, comprising 40 frames each, were acquired in counting mode. The range of defocus values spanned from 0.7 to 1.7 μm. Detailed information about the Number of movies and the total electron dose for each dataset is presented in Table 4.26.

Table 4.26 Variability in data acquisition parameters across collected datasets.

Datasets	Image stacks	Total electron dose (electrons/Å²)	WARP-picked particles
Dataset1	19047	39.6	1,004,005
Dataset2	16289	43.6	960,668
Dataset3	11644	43.2	1,344,141
Dataset4	16169	43.2	1,566,773
Dataset5	15614	44.4	1,129,934

The preprocessing steps for all five datasets were consistent with those applied to the hTFIIIC-DNA complex, as detailed in section 4.9.2. Notably, Dataset1 underwent preprocessing using RELION 3.1.3, whereas Datasets 2 through 5 were processed with RELION 4.0 [143]. The versions of cryoSPARC used ranged from 3.3.2 to 4.4, concurrent with the processing period of the datasets. For all datasets, the initial particles identified by WARP were extracted with a box size of 480 px.

Specifically, for the τB-DNA subcomplex processing, the WARP-picked particles in each dataset were classified involving 2-3 rounds of heterogeneous refinement in cryoSPARC. This process included one volume derived from an ab-initio job, based on selected 2D classes resembling the τB-DNA complex, and

2-3 'junk' classes formed using a subset of randomly chosen particles (refer to Table 4.26 for initial WARP-picked particles). A non-uniform refinement was applied to the best map from the final round of heterogeneous refinement for each dataset, achieving resolutions between 2.98 Å and 3.25 Å. Subsequently, these particles, with refined Euler angles, were re-imported into RELION for re-extraction using the initial box size of 480 px and then returned to cryoSPARC for final processing. The contribution of particles from each dataset to the τB subcomplex prior to the final classification step is detailed in Table 4.27. At this stage the total amount of particles from all five datasets reached 1,280,178. These particles were further classified through two rounds of heterogeneous refinement in cryoSPARC. The highest-quality map of the τB-DNA complex, comprising 833,266 particles, was selected for a final non-uniform refinement, achieving a resolution of 2.67 Å.

Table 4.27. Particle count from each dataset to the final τB-DNA complex classification before dataset merging.

Datasets	τB-DNA particles
Dataset1	191,618
Dataset2	281,077
Dataset3	326,061
Dataset4	226,854
Dataset5	254,568

For the τA-DNA complex, each dataset was processed through an identical workflow. After initial particle picking using WARP (referenced in Table 4.26), these particles were subjected to 2D classification. Classes displaying distinct τA-DNA complex features were chosen for the first round of TOPAZ training. These classes were then re-imported to RELION, where particles were extracted with a new box size of 300px, specifically for TOPAZ training. Post-training and picking, the TOPAZ-picked particles were extracted at a 320 px box size and imported to cryoSPARC. In cryoSPARC, these particles were classified by heterogeneous refinement, followed by 2D classification to remove clearly junk particles. These classified particle sets were then used for a second round of TOPAZ training, a similar particle re-extraction was applied in comparison with the first TOPAZ training. After the second training and picking step, the particles were again classified through heterogeneous refinement and 2D classification in cryoSPARC. At this stage, only junk classes were discarded. The remaining particles from each dataset were then combined for subsequent analyses. Table 4.28 details the particle counts contributed by each dataset to this combined set. A total of 739 929 particles were processed through two iterations of heterogeneous refinement. Following this, the class with the best map, containing 114 621 particles, was selected for non-uniform refinement, resulting in a resolution of 3.73 Å for the τA-DNA complex.

Table 4.28 Particle count from each dataset before merging particles for the final round of τA-DNA classification.

Datasets	τA-DNA particles
Dataset1	178,188
Dataset2	94,519
Dataset3	139,899
Dataset4	177,894
Dataset5	149,429

Despite the addition of various transcription factors during sample preparation (refer to Table 4.25), the final reconstructions of the τA-DNA and τB-DNA complexes revealed no additional density attributable to these added transcription factors.

4.10.3 Model building, refinement and validation

A methodology similar to that used for the hTFIIIC and hTFIIIC-DNA complexes was applied for model building. Initially, two AlphaFold Multimer predictions were performed: one incorporating the C-terminus of τ138 (amino acids 635 to 1060), τ131, τ95, and τ55 subunits to form the τA complex, and the other including the N-terminus of τ138 (amino acids 1 to 668), τ91, and τ60 for the τB complex. These structural predictions were subsequently rigidly docked into their respective densities using ChimeraX. The .pkl files generated from these predictions were converted into .json files, featuring AlphaFold prediction scores, utilizing a script from http://www.subtiwiki.uni-goettingen.de/v4/paeViewerDemo. ISOLDE [24] was then used to refine these initial AlphaFold models of τA and τB into their densities. This refinement was particularly crucial for fitting secondary structures within areas of lower local resolution (ranging between 4.5 to 6 Å), specially for the τA map. Several rounds of ISOLDE were implemented to fix rotamer and Ramachandran outliers, followed by iterative refinement using PHENIX. In the τB-DNA complex, due to its higher resolution, a B-DNA model corresponding to the first 40 nucleotides of the downstream region of the DNA oligo was confidently positioned. Conversely, for the τA-DNA complex, the adjacent upstream 45 base pairs were modeled into the density. However, the low local resolution in the τA-DNA complex introduces some ambiguity regarding the precise orientation of the DNA.

IV. Appendix

5. Appendix

Table A1. Cryo-EM data collection, refinement and validation statistics of the human TFIIIC structures and maps. Modified table taken from [66]

	τB-DNA DNA sample Monomer 1 map (EMD-16713) (PDB 8CLI)	τB-DNA dimer map DNA sample Composite map (EMD-16714, EMD-17446, EMD-17447) (PDB 8CLJ)	τA map DNA sample (EMD-16715) (PDB 8CLK)	τB dimer map No DNA sample (EMD-16717) (PDB 8CLL)
Data collection and processing				
Magnification	105,000	105,000	105,000	105,00
Voltage (kV)	300	300	300	300
Electron exposure (e–/Å^2)	42.8	42.8	42.8	39.6
Defocus range (μm)	0.7-1.7	0.7-1.7	0.7-1.7	0.7-1.7
Pixel size (Å)	0.822	0.822	0.822	0.822
Symmetry imposed	C1	C1	C1	C1
Initial particle images (no.)	762,590	762,590	365,063	754,951
Final particle images (no.)	35,379	35,379	55,079	125,600
Map resolution (Å)	3.2	3.2	3.5	3.4
FSC threshold	0.143	0.143	0.143	0.143
Map resolution range (Å)	3.1-9.9	n/a	3.4-13.4	3.2-10.9
Refinement				
Initial model used (Alphafold structure prediction)	AF-Q12789-F1, AF-Q8WUA4-F1, AF-Q9UKN8-F1	AF-Q12789-F1, AF-Q8WUA4-F1, AF-Q9UKN8-F1	AF-Q9Y5Q9-F1, AF-Q9Y5Q8-F1, AF-Q969F1-F1	AF-Q12789-F1, AF-Q8WUA4-F1, AF-Q9UKN8-F1
Model resolution (Å)	3.3	3.3	3.7	3.4
FSC threshold	0.5	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	-75.9	-	-114.5	-110.6
Model composition				
Non-hydrogen atoms	15955	31912	10800	25288
Protein / nucleotide residues	1810 / 70	3620 / 140	1344 / 0	3152 / 0
Ligands	2x Zn	4x Zn	-	4x Zn
***B* factors (Å^2)**				
Protein (min/max/mean)	22.5/314.4 /95.5	12.4/303.3 /93.0	5.1/259.8 /117.6	13.0/306.9/115.1
Ligand (min/max/mean)	148.7/165.1/156.9	125.9/167.6/149.3	-	188.7/250.6/222.1
Nucleotide (min/max/mean)	5.9/207.1/104.1		-	-

		2.4/301.9 /113.4		
R.m.s. deviations				
Bond lengths (Å)	0.004	0.004	0.005	0.003
Bond angles (°)	0.716	0.753	0.754	0.662
Validation				
MolProbity score	1.13	1.11	1.31	1.03
Clashscore	2.62	2.74	4.17	2.24
Poor rotamers (%)	0.00	0.00	0.00	0.00
Ramachandran plot				
Favored (%)	97.6	97.8	97.4	97.9
Allowed (%)	2.4	2.2	2.6	2.1
Disallowed (%)	0.0	0.0	0.0	0.0
Rama-Z score				
Whole	-0.73(0.18)	-0.77(0.13)	-1.42(0.21)	-0.75(0.14)
Helix	-0.89(0.19)	-0.86(0.13)	-1.12(0.17)	-0.76(0.15)
Sheet	0.25(0.24)	0.23(0.17)	-0.30(0.45)	0.05(0.18)
Loop	-0.46(0.21)	-0.53(0.15)	-0.63(0.26)	-0.52(0.15)

Table A2. Cryo-EM data collection and refinement for human τA without DNA. Modified table taken from [66]

	τA map No DNA sample (EMD-16716)
Data collection and processing	
Magnification	105,00
Voltage (kV)	300
Electron exposure (e–/Å^2)	39.6
Defocus range (μm)	0.7-1.7
Pixel size (Å)	0.822
Symmetry imposed	C1
Initial particle images (no.)	748,471
Final particle images (no.)	94,103
Map resolution (Å)	3.8
FSC threshold	0.143
Map resolution range (Å)	3.7-14

Table A3. Cryo-EM data collection, refinement and validation statistics of the yeast TFIIIC structures

	τB-DNA monomer	τA-DNA
Data collection and processing		
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	39.6, 43.6, 43.2, 43.2, 44.4	39.6, 43.6, 43.2, 43.2, 44.4
Defocus range (μm)	0.7-1.7	0.7-1.7
Pixel size (Å)	0.822	0.822
Symmetry imposed	C1	C1
Initial particle images (no.)	1,280,178	739,929
Final particle images (no.)	833,266	114,321
Map resolution (Å)	2.67	3.73
FSC threshold	0.143	0.143
Map resolution range (Å)	2.6-9.9	3.6-13.8
Refinement		
Initial model used	Alphafold multimer prediction	Alphafold multimer prediction
Model resolution (Å)	2.75 and 2.5	7.5 and 3.4
FSC threshold	0.5	0.5
FSC threshold	0.143	0.143
Map sharpening *B* factor (Å^2)	-111.4	-159.7
Model composition		
Non-hydrogen atoms	15002	12264
Protein/Nucleotide residues	1667 / 80	1283 / 90
Ligands	-	-
***B* factors (Å^2)**		
Protein (min/max/mean)	10.65/176.58	34.91/330.86
Ligand (min/max/mean)	/61.73	/158.47
Nucleotide (min/max/mean)	- 00/112.8/41.76	- 128.86/570.04 /263.35
R.m.s. deviations		
Bond lengths (Å)	0.004	0.005
Bond angles (°)	0.730	0.867
Validation		
MolProbity score	1.03	1.92
Clashscore	1.77	9.64
Poor rotamers (%)	0.00	0.09
Ramachandran plot		

Favored (%)	97.53	93.7
Allowed (%)	2.47	6.15
Disallowed (%)	0.0	0.16
Rama-Z score		
Whole	-0.47(0.19)	-2.16(0.21)
Helix	-1.02(0.22)	-1.54(0.17)
Sheet	0.35(0.24)	-0.34(0.46)
Loop	-0.18(0.21)	-1.41(0.24)

Abbreviations

Å: Angstroms

bp: Base pair

Cryo-EM: cryo-electron microscopy

CTF: Contrast transfer function

CV: column volume

D: Dalton

DBD: DNA binding domain

DTT: Dithiothreitol

ddH$_2$O: double water distilled !

EMSA: Electrophoretic mobility shift assay

eWH: extended winged helix domain

GFP: Green Fluorescence Protein

GA: Glutaraldehyde

hTFIIIC: human transcription factor IIIC

IRB: Isothermal Reaction Buffer

ITC: Isothermal titration calorimetry

LMNG: Lauryl Maltose Neopentyl Glycol

min: minute

MP: mass photometry

MW: molecular weight

OG: Octyl-beta-Glucoside

OSC: Oxygen Scavenger

PBS: Phosphate buffered saline

PCR: polymerase chain reactions

px: pixel

RT: Room temperature

s: second

SMFM: single-molecule fluorescence microscopy

yTFIIIC: yeast transcription factor IIIC

TBE: Tris Borate EDTA buffer

T_m: melting temperature

V: Volt

WH: winged-helix domain

6. References

[1]   F. Werner, "Structural evolution of multisubunit RNA polymerases," *Trends Microbiol.*, vol. 16, no. 6, pp. 247–250, 2008, doi: 10.1016/j.tim.2008.03.008.

[2]   F. Werner and D. Grohmann, "Evolution of multisubunit RNA polymerases in the three domains of life," *Nat. Rev. Microbiol.*, vol. 9, no. 2, pp. 85–98, 2011, doi: 10.1038/nrmicro2507.

[3]   F. Werner, "Structure and function of archaeal RNA polymerases," *Mol. Microbiol.*, vol. 65, no. 6, pp. 1395–1404, 2007, doi: 10.1111/j.1365-2958.2007.05876.x.

[4]   J. Hurwitz, "The discovery of RNA polymerase," *J. Biol. Chem.*, vol. 280, no. 52, pp. 42477–42485, 2005, doi: 10.1074/jbc.X500006200.

[5]   Darst S. A., Kubalek E. W., and Kornberg R.D., "Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography," *Nature*, vol. 340, no. August, pp. 730–732, 1989.

[6]   K. S. Murakami, "X-ray crystal structure of escherichia coli RNA polymerase σ70 holoenzyme," *J. Biol. Chem.*, vol. 288, no. 13, pp. 9126–9134, 2013, doi: 10.1074/jbc.M112.430900.

[7]   C. Sutherland and K. S. Murakami, " An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia coli RNA Polymerase ," *EcoSal Plus*, vol. 8, no. 1, pp. 1–14, 2018, doi: 10.1128/ecosalplus.esp-0004-2018.

[8]   E. A. Campbell *et al.*, "Structure of the bacterial RNA polymerase promoter specificity σ subunit," *Molecular Cell*, vol. 9, no. 3. pp. 527–539, 2002. doi: 10.1016/S1097-2765(02)00470-7.

[9]   K. S. Murakami and S. A. Darst, "Bacterial RNA polymerases: The wholo story," *Curr. Opin. Struct. Biol.*, vol. 13, no. 1, pp. 31–39, 2003, doi: 10.1016/S0959-440X(02)00005-2.

[10]   Y. Korkhin *et al.*, "Evolution of complex RNA polymerases: The complete archaeal RNA polymerase structure," *PLoS Biol.*, vol. 7, no. 5, 2009, doi: 10.1371/journal.pbio.1000102.

[11]   S. D. Bell, C. Jaxel, M. Nadal, P. F. Kosa, and S. P. Jackson, "Temperature, template topology, and factor requirements of archaeal transcription," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 26, pp. 15218–15222, 1998, doi: 10.1073/pnas.95.26.15218.

[12]   M. Ouhammouch, "Transcriptional regulation in Archaea," *Curr. Opin. Genet. Dev.*, vol. 14, no. 2, pp. 133–138, 2004, doi: 10.1016/j.gde.2004.01.002.

[13]   S. A. Qureshi and S. P. Jackson, "Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength," *Mol. Cell*, vol. 1, no. 3, pp. 389–400, 1998, doi: 10.1016/S1097-2765(00)80039-8.

[14]   O. Littlefield, Y. Korkhin, and P. B. Sigler, "The structural basis for the oriented assembly of a TBP/TFB/promoter complex," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 96, no. 24, pp. 13668–13673, 1999, doi: 10.1073/pnas.96.24.13668.

[15]   P. Cramer *et al.*, "Structure of eukaryotic RNA polymerases," *Annu. Rev. Biophys.*, vol. 37, pp. 337–352, 2008, doi: 10.1146/annurev.biophys.37.032807.130008.

[16]   C. Fernández-Tornero *et al.*, "Crystal structure of the 14-subunit RNA polymerase I," *Nature*, vol. 502, no. 7473, pp. 644–649, 2013, doi: 10.1038/nature12636.

[17]   A. D. Misiaszek *et al.*, "Cryo-EM structures of human RNA polymerase I," *Nat. Struct. Mol. Biol.*, vol. 28, no. 12, pp. 997–1008, 2021, doi: 10.1038/s41594-021-00693-4.

[18]   C. Bernecky, F. Herzog, W. Baumeister, J. M. Plitzko, and P. Cramer, "Structure of transcribing mammalian RNA polymerase II," *Nature*, vol. 529, no. 7587, pp. 551–554, 2016, doi: 10.1038/nature16482.

[19]   P. Cramer, D. A. Bushnell, and R. D. Kornberg, "Structural basis of transcription: RNA polymerase II at 2.8 ångstrom resolution," *Science (80-.).*, vol. 292, no. 5523, pp. 1863–1876, 2001, doi: 10.1126/science.1059493.

[20]   M. Girbig, A. D. Misiaszek, and C. W. Müller, "Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases," *Nat. Rev. Mol. Cell Biol.*, vol. 23, no. 9, pp. 603–622, Sep. 2022, doi: 10.1038/s41580-022-00476-9.

[21]   M. Cieśla, E. Skowronek, and M. Boguta, "Function of TFIIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription," *Nucleic Acids Res.*, vol. 46, no. 18, pp. 9444–9455, 2018, doi: 10.1093/nar/gky656.

[22]   R. Ferrari, C. Rivetti, J. Acker, and G. Dieci, "Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 37, pp. 13442–13447, 2004, doi: 10.1073/pnas.0403851101.

[23]   R. Arrebola *et al.*, "τ91, an Essential Subunit of Yeast Transcription Factor IIIC, Cooperates with τ138 in DNA Binding," *Mol. Cell. Biol.*, vol. 18, no. 1, pp. 1–9, 1998, doi: 10.1128/mcb.18.1.1.

[24]   D. Canella, V. Praz, J. H. Reina, P. Cousin, and N. Hernandez, "Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells," *Genome Res.*, vol. 20, no. 6, pp. 710–721, 2010, doi: 10.1101/gr.101337.109.

[25]   G. R. Kunkel, "RNA polymerase III transcription of genes that lack internal control regions," *BBA - Gene Struct. Expr.*, vol. 1088, no. 1, pp. 1–9, 1991, doi: 10.1016/0167-4781(91)90146-D.

[26]   E. Guffanti *et al.*, "A minimal promoter for TFIIIC-dependent in vitro transcription of snoRNA and tRNA genes by RNA polymerase III," *J. Biol. Chem.*, vol. 281, no. 33, pp. 23945–23957, 2006, doi: 10.1074/jbc.M513814200.

[27]   D. R. Engelke, S. Y. Ng, B. S. Shastry, and R. G. Roeder, "Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes," *Cell*, vol. 19, no. 3, pp. 717–728, 1980, doi: 10.1016/S0092-8674(80)80048-1.

[28]   E. P. Geiduschek and G. P. Tocchini-Valentini, "Transcription by RNA polymerase III," *Annu. Rev. Biochem.*, vol. 57, pp. 873–914, 1988, doi: 10.1146/annurev.bi.57.070188.004301.

[29]   J. S. Hanas, D. J. Hazuda, D. F. Bogenhagen, F. Y. Wu, and C. W. Wu, "Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene.," *J. Biol. Chem.*, vol. 258, no. 23, pp. 14120–14125, 1983, doi: 10.1016/s0021-9258(17)43831-2.

[30] S. Camier, A. M. Dechampesme, and A. Sentenac, "The only essential function of TFIIIA in yeast is the transcription of 5S rRNA genes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 20, pp. 9338–9342, 1995, doi: 10.1073/pnas.92.20.9338.

[31] A. B. Lassar, P. L. Martin, and R. G. Roeder, "Transcription of Class III Genes: Formation of Preinitiation Complexes," *Science (80-.).*, vol. 222, no. 4625, pp. 740–748, Nov. 1983, doi: 10.1126/science.6356356.

[32] G. Galli, H. Hofstetter, and M. L. Birnstiel, "Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements," *Nature*, vol. 294, pp. 626–631, 1981.

[33] R. E. Baker, S. Camier, A. Sentenac, and B. D. Hall, "Gene size differentially affects the binding of yeast transcription factor tau to two intragenic regions.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 84, no. 24, pp. 8768–8772, 1987, doi: 10.1073/pnas.84.24.8768.

[34] A. Ruet, S. Camier, W. Smagowicz, A. Sentenac, and P. Fromageot, "Isolation of a class C transcription factor which forms a stable complex with tRNA genes," *EMBO J.*, vol. 3, no. 2, pp. 343–350, Feb. 1984, doi: 10.1002/j.1460-2075.1984.tb01809.x.

[35] R. E. Baker, O. Gabrielsen, and B. D. Hall, "Effects of tRNATyr point mutations on the binding of yeast RNA polymerase III transcription factor C.," *J. Biol. Chem.*, vol. 261, no. 12, pp. 5275–5282, Apr. 1986, doi: 10.1016/S0021-9258(19)57210-6.

[36] N. M. I. Taylor, F. Baudin, G. Von Scheven, and C. W. Müller, "RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74," *Nucleic Acids Res.*, vol. 41, no. 19, pp. 9183–9196, 2013, doi: 10.1093/nar/gkt664.

[37] M. K. Vorländer *et al.*, "Structure of the TFIIIC subcomplex τ A provides insights into RNA polymerase III pre-initiation complex formation," *Nature Communications*. pp. 1–12, 2020. [Online]. Available: http://dx.doi.org/10.1038/s41467-020-18707-y

[38] G. A. Kassavetis, D. L. Riggs, R. Negri, L. H. Nguyen, and E. P. Geiduschek, "Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes," *Mol. Cell. Biol.*, vol. 9, no. 171, pp. 2551–2566, Jun. 1989, doi: 10.1128/mcb.9.6.2551-2566.1989.

[39] L. Schramm and N. Hernandez, "Recruitment of RNA polymerase III to its target promoters," *Genes Dev.*, vol. 16, no. 20, pp. 2593–2620, 2002, doi: 10.1101/gad.1018902.

[40] A. Krol and L. De Biochimie, "Xenopus tropicals U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes," *Nucleic Acids Res.*, vol. 15, no. 6, pp. 2463–2478, 1987.

[41] R. Waldschmidt, I. Wanandi, and K. H. Seifart, "Identification of transcription factors required for the expression of mammalian U6 genes in vitro.," *EMBO J.*, vol. 10, no. 9, pp. 2595–2603, 1991, doi: 10.1002/j.1460-2075.1991.tb07801.x.

[42] S. Murphy, J.-B. Yoon, T. Gerster, and R. G. Roeder, "Oct-1 and Oct-2 Potentiate Functional Interactions of a Transcription Factor with the Proximal Sequence Element of Small Nuclear RNA Genes," *Mol. Cell. Biol.*, vol. 12, no. 7, pp. 3247–3261, 1992, doi: 10.1128/mcb.12.7.3247-3261.1992.

[43]   C. L. Sadowski, R. W. Henry, R. Kobayashi, and N. Hernandez, "The SNAP45 subunit of the small nuclear RNA (snRNA) activating protein complex is required for RNA polymerase II and III snRNA gene transcription and interacts with the TATA box binding protein," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 9, pp. 4289–4293, 1996, doi: 10.1073/pnas.93.9.4289.

[44]   S. Camier, O. Gabrielsen, R. Baker, and A. Sentenac, "A split binding site for transcription factor tau on the tRNA3Glu gene," *EMBO J.*, vol. 4, no. 2, pp. 491–500, 1985, doi: 10.1002/j.1460-2075.1985.tb03655.x.

[45]   R. E. Baker, O. Gabrielsen, and B. D. Hall, "Effects of tRNATyr point mutations on the binding of yeast RNA polymerase III transcription factor C.," *J. Biol. Chem.*, vol. 261, no. 12, pp. 5275–5282, Apr. 1986, doi: 10.1016/S0021-9258(19)57210-6.

[46]   N. Marzouki, S. Camier, A. Ruet, A. Moenne, and A. Sentenac, "Selective proteolysis defines two DNA binding domains in yeast transcription factor τ," *Nature*, vol. 323, pp. 176–178, Sep. 1986, doi: 10.1038/323176a0.

[47]   P. Schultz, N. Marzouki, C. Marck, A. Ruet, P. Oudet, and A. Sentenac, "The two DNA-binding domains of yeast transcription factor τ as observed by scanning transmission electron microscopy," *EMBO J.*, vol. 8, no. 12, pp. 3815–3824, 1989, doi: 10.1002/j.1460-2075.1989.tb08559.x.

[48]   O. S. Gabrielsen, N. Marzouki, A. Ruet, A. Sentenac, and P. Fromageot, "Two polypeptide chains in yeast transcription factor τ interact with DNA," *J. Biol. Chem.*, vol. 264, no. 13, pp. 7505–7511, 1989, doi: 10.1016/s0021-9258(18)83263-x.

[49]   B. Bartholomew, G. A. Kassavetis, B. R. Braun, and E. P. Geiduschek, "The subunit structure of Saccharomyces cerevisiae transcription factor IIIC probed with a novel photocrosslinking reagent.," *EMBO J.*, vol. 9, no. 7, pp. 2197–2205, 1990, doi: 10.1002/j.1460-2075.1990.tb07389.x.

[50]   R. N. Swanson *et al.*, "Isolation of TFC1, a gene encoding one of two DNA-binding subunits of yeast transcription factor τ (TFIIIC)," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 11, pp. 4887–4891, 1991, doi: 10.1073/pnas.88.11.4887.

[51]   M. C. Parsons and P. A. Weil, "Cloning of TFC1, the Saccharomyces cerevisiae gene encoding the 95-kDa subunit of transcription factor TFIIIC," *J. Biol. Chem.*, vol. 267, no. 5, pp. 2894–2901, 1992, doi: 10.1016/s0021-9258(19)50670-6.

[52]   O. Lefebvre *et al.*, "TFC3: gene encoding the B-block binding subunit of the yeast transcription factor IIIC.," *Proc. Natl. Acad. Sci.*, vol. 89, no. 21, pp. 10512–10516, Nov. 1992, doi: 10.1073/pnas.89.21.10512.

[53]   C. Marck *et al.*, "The TFIIIB-assembling subunit of yeast transcription factor TFIIIC has both tetratricopeptide repeats and basic helix-loop-helix motifs," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 90, no. 9, pp. 4027–4031, 1993, doi: 10.1073/pnas.90.9.4027.

[54]   N. Manaud *et al.*, "A Chimeric Subunit of Yeast Transcription Factor IIIC Forms a Subcomplex with τ95," *Mol. Cell. Biol.*, vol. 18, no. 6, pp. 3191–3200, 1998, doi: 10.1128/mcb.18.6.3191.

[55]   E. Deprez, R. Arrebola, C. Conesa, and A. Sentenac, "A Subunit of Yeast TFIIIC Participates in the Recruitment of TATA-Binding Protein," *Mol. Cell. Biol.*, vol. 19, no. 12, pp. 8042–8051, 1999, doi:

10.1128/mcb.19.12.8042.

[56]   C. Conesa, R. N. Swanson, P. Schultz, P. Oudet, and A. Sentenac, "On the subunit composition, stoichiometry, and phosphorylation of the yeast transcription factor TFIIIC/τ," *J. Biol. Chem.*, vol. 268, no. 24, pp. 18047–18052, 1993, doi: 10.1016/s0021-9258(17)46809-8.

[57]   B. R. Braun, B. Bartholomew, G. A. Kassavetis, and E. P. Geiduschek, "Topography of transcription factor complexes on the Saccharomyces cerevisiae 5 S RNA gene," *J. Mol. Biol.*, vol. 228, no. 4, pp. 1063–1077, 1992, doi: 10.1016/0022-2836(92)90315-B.

[58]   C. Ducrot, O. Lefebvre, E. Landrieux, J. Guirouilh-Barbat, A. Sentenac, and J. Acker, "Reconstitution of the yeast RNA polymerase III transcription system with all recombinant factors," *J. Biol. Chem.*, vol. 281, no. 17, pp. 11685–11692, 2006, doi: 10.1074/jbc.M600101200.

[59]   A. Mylona *et al.*, "Structure of the τ60/Δτ91 Subcomplex of Yeast Transcription Factor IIIC: Insights into Preinitiation Complex Assembly," *Mol. Cell*, vol. 24, no. 2, pp. 221–232, 2006, doi: 10.1016/j.molcel.2006.08.013.

[60]   G. Male *et al.*, "Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly," *Nat. Commun.*, vol. 6, no. 1, p. 7387, Nov. 2015, [Online]. Available: http://www.nature.com/articles/ncomms8387

[61]   N. M. I. Taylor *et al.*, "Structural and functional characterization of a phosphatase domain within yeast general transcription factor IIIC," *J. Biol. Chem.*, vol. 288, no. 21, pp. 15110–15120, 2013, doi: 10.1074/jbc.M112.427856.

[62]   N. M. I. Taylor, F. Baudin, G. von Scheven, and C. W. Müller, "RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74," *Nucleic Acids Res.*, vol. 41, no. 19, pp. 9183–9196, Oct. 2013, doi: 10.1093/nar/gkt664.

[63]   W. Kühlbrandt, "The Resolution Revolution," *Science (80-.).*, vol. 343, no. 6178, pp. 1443–1444, 2014, doi: 10.1126/science.1251652.

[64]   J. Hanske, Y. Sadian, and C. W. Müller, "The cryo-EM resolution revolution and transcription complexes," *Curr. Opin. Struct. Biol.*, vol. 52, pp. 8–15, 2018, doi: 10.1016/j.sbi.2018.07.002.

[65]   A. Talyzina *et al.*, "Structural basis of TFIIIC-dependent RNA polymerase III transcription initiation," *Mol. Cell*, vol. 83, no. 15, pp. 2641-2652.e7, 2023, doi: 10.1016/j.molcel.2023.06.015.

[66]   W. Seifert-Davila, M. Girbig, L. Hauptmann, T. Hoffmann, S. Eustermann, and C. W. Muller, "Structural insights into human TFIIIC promoter recognition," *Sci. Adv.*, vol. 9, no. 27, pp. 1–11, 2023, doi: 10.1126/sciadv.adh2019.

[67]   S. K. Yoshinaga, P. A. Boulanger, and A. J. Berk, "Resolution of human transcription factor TFIIIC into two functional components," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 84, no. 11, pp. 3585–3589, 1987, doi: 10.1073/pnas.84.11.3585.

[68]   N. Dean and A. J. Berk, "Separation of TFIIIC into two functional components by sequence specific DNA affuinity chromatography," *Nucleic Acids Res.*, vol. 15, no. 23, pp. 9895–9907, 1987.

[69]   J. A. Cromlish and R. G. Roeder, "Human transcription factor IIIC (TFIIIC)," *J. Biol. Chem.*, vol. 264,

no. 30, pp. 18100–18109, 1989, doi: 10.1016/s0021-9258(19)84683-5.

[70]   S. K. Yoshinaga, N. D. L'Etoile, and A. J. Berk, "Purification and characterization of transcription factor IIIC2," *J. Biol. Chem.*, vol. 264, no. 18, pp. 10726–10731, 1989, doi: 10.1002/mas.21369.

[71]   R. Kovelman and R. G. Roeder, "Purification and characterization of two forms of human transcription factor IIIC," *J. Biol. Chem.*, vol. 267, no. 34, pp. 24446–24456, 1992, doi: 10.1016/s0021-9258(18)35787-9.

[72]   N. D. L'Etoile, M. L. Fahnestock, Y. Shen, R. Aebersold, and A. J. Berk, "Human transcription factor IIIC box B binding subunit," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 91, no. 5, pp. 1652–1656, 1994, doi: 10.1073/pnas.91.5.1652.

[73]   E. Sinn, Z. Wang, R. Kovelman, and R. G. Roeder, "Cloning and characterization of a TFIIIC2 subunit (TFIIICβ) whose presence correlates with activation of RNA polymerase III-mediated transcription by adenovirus E1A expression and serum factors," *Genes Dev.*, vol. 9, no. 6, pp. 675–685, 1995, doi: 10.1101/gad.9.6.675.

[74]   Y.-J. Hsieh, Z. Wang, R. Kovelman, and R. G. Roeder, "Cloning and Characterization of Two Evolutionarily Conserved Subunits (TFIIIC102 and TFIIIC63) of Human TFIIIC and Their Involvement in Functional Interactions with TFIIIB and RNA Polymerase III," *Mol. Cell. Biol.*, vol. 19, no. 7, pp. 4944–4952, 1999, doi: 10.1128/mcb.19.7.4944.

[75]   Y.-J. Hsieh, T. K. Kundu, Z. Wang, R. Kovelman, and R. G. Roeder, "The TFIIIC90 Subunit of TFIIIC Interacts with Multiple Components of the RNA Polymerase III Machinery and Contains a Histone-Specific Acetyltransferase Activity," *Mol. Cell. Biol.*, vol. 19, no. 11, pp. 7697–7704, 1999, doi: 10.1128/mcb.19.11.7697.

[76]   H. Dumay-Odelot *et al.*, "Identification, molecular cloning, and characterization of the sixth subunit of human transcription factor TFIIIC," *J. Biol. Chem.*, vol. 282, no. 23, pp. 17179–17189, 2007, doi: 10.1074/jbc.M611542200.

[77]   N. A. Hoffmann, A. J. Jakobi, M. K. Vorländer, C. Sachse, and C. W. Müller, "Transcribing RNA polymerase III observed by electron cryomicroscopy," *FEBS J.*, vol. 283, pp. 2811–2819, 2016, doi: 10.1111/febs.13732.

[78]   S. Weser *et al.*, "Transcription factor (TF)-like nuclear regulator, the 250-kDa form of homo sapiens TFIIIB", is an essential component of human TFIIIC1 activity," *Journal of Biological Chemistry*, vol. 279, no. 26. pp. 27022–27029, 2004. doi: 10.1074/jbc.M312790200.

[79]   M. K. Vorländer *et al.*, "Structure of the TFIIIC subcomplex τA provides insights into RNA polymerase III pre-initiation complex formation," *Nat. Commun.*, vol. 11, no. 1, p. 4905, Dec. 2020, doi: 10.1038/s41467-020-18707-y.

[80]   J. Gouge *et al.*, "Redox Signaling by the RNA Polymerase III TFIIB-Related Factor Brf2," *Cell*, vol. 163, no. 6, pp. 1375–1387, Dec. 2015, doi: 10.1016/j.cell.2015.11.005.

[81]   G. A. Kassavetis, B. R. Braun, L. H. Nguyen, and G. E. Peter, "S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors," *Cell*, vol. 60, no. 2, pp. 235–245, Jan. 1990, doi: 10.1016/0092-8674(90)90739-2.

[82]   J. Gouge *et al.*, "Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation," *Nat. Commun.*, vol. 8, no. 1, p. 130, Jul. 2017, doi: 10.1038/s41467-017-00126-1.

[83]   G. Abascal-Palacios, E. P. Ramsay, F. Beuron, E. Morris, and A. Vannini, "Structural basis of RNA polymerase III transcription initiation," *Nature*, vol. 553, no. 7688, pp. 301–306, 2018, doi: 10.1038/nature25441.

[84]   M. K. Vorländer, H. Khatter, R. Wetzel, W. J. H. Hagen, and C. W. Müller, "Molecular mechanism of promoter opening by RNA polymerase III," *Nature*, vol. 553, no. 7688, pp. 295–300, 2018, doi: 10.1038/nature25440.

[85]   D. N. Roberts, A. J. Stewart, J. T. Huff, and B. R. Cairns, "The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 25, pp. 14695–14700, 2003, doi: 10.1073/pnas.2435566100.

[86]   O. Harismendy *et al.*, "Genome-wide location of yeast RNA polymerase III transcription machinery," *EMBO J.*, vol. 22, no. 18, pp. 4738–4747, 2003, doi: 10.1093/emboj/cdg466.

[87]   Z. Moqtaderi and K. Struhl, "Genome-Wide Occupancy Profile of the RNA Polymerase III Machinery in Saccharomyces cerevisiae Reveals Loci with Incomplete Transcription Complexes," *Mol. Cell. Biol.*, vol. 24, no. 10, pp. 4118–4127, 2004, doi: 10.1128/mcb.24.10.4118-4127.2004.

[88]   K. ichi Noma, H. P. Cam, R. J. Maraia, and S. I. S. Grewal, "A Role for TFIIIC Transcription Factor Complex in Genome Organization," *Cell*, vol. 125, no. 5, pp. 859–872, 2006, doi: 10.1016/j.cell.2006.04.028.

[89]   S. I. Hiraga, S. Botsios, D. Donze, and A. D. Donaldson, "TFIIIC localizes budding yeast ETC sites to the nuclear periphery," *Mol. Biol. Cell*, vol. 23, no. 14, pp. 2741–2754, 2012, doi: 10.1091/mbc.E11-04-0365.

[90]   Z. Moqtaderi *et al.*, "Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells," *Nat. Struct. Mol. Biol.*, vol. 17, no. 5, pp. 635–640, 2010, doi: 10.1038/nsmb.1794.

[91]   L. Carrire *et al.*, "Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells," *Nucleic Acids Res.*, vol. 40, no. 1, pp. 270–283, 2012, doi: 10.1093/nar/gkr737.

[92]   K. C. Yuen, B. D. Slaughter, and J. L. Gerton, "Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters," *Sci. Adv.*, vol. 3, no. 6, 2017, doi: 10.1126/sciadv.1700191.

[93]   R. Ferrari *et al.*, "TFIIIC Binding to Alu Elements Controls Gene Expression via Chromatin Looping and Histone Acetylation," *Mol. Cell*, vol. 77, no. 3, pp. 475-487.e11, 2020, doi: 10.1016/j.molcel.2019.10.020.

[94]   L. de L. Cucalon *et al.*, "An RNA Polymerase III General Transcription Factor Engages in Cell Type-Specific Chromatin Looping," *Int. J. Mol. Sci.*, vol. 23, no. 4, 2022, doi: 10.3390/ijms23042260.

[95]   R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing, "Model for

the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy," *J. Mol. Biol.*, vol. 213, no. 4, pp. 899–929, 1990, [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0022283605802712

[96]   E. Nogales, "Profile of Joachim Frank, Richard Henderson, and Jacques Dubochet, 2017 Nobel Laureates in Chemistry," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 3, pp. 441–444, 2018, doi: 10.1073/pnas.1718898114.

[97]   B. E. Bammesa, R. H. Rochata, J. Jakanab, D.-H. Chenb, and W. Chiua, "Direct electron detection yields cryo-EM reconstructions at resolutions beyond ¾ Nyquist frequency," *J. Struct. Biol.*, vol. 177, no. 3, pp. 589–601, 2012, doi: 10.1016/j.jsb.2012.01.008.Direct.

[98]   S. H. W. Scheres, "RELION: Implementation of a Bayesian approach to cryo-EM structure determination," *J. Struct. Biol.*, vol. 180, no. 3, pp. 519–530, 2012, doi: 10.1016/j.jsb.2012.09.006.

[99]   L. A. Passmore and C. J. Russo, "Specimen preparation for high-resolution cryo-EM," *Methods Enzymol.*, no. 579, pp. 51–86, 2016, doi: 10.1016/bs.mie.2016.04.011.Specimen.

[100]   G. Weissenberger, R. J. M. Henderikx, and P. J. Peters, "Understanding the invisible hands of sample preparation for cryo-EM," *Nat. Methods*, vol. 18, no. 5, pp. 463–471, 2021, doi: 10.1038/s41592-021-01130-6.

[101]   M. Armstrong, B. G. Han, S. Gomez, J. Turner, D. A. Fletcher, and R. M. Glaeser, "Microscale Fluid Behavior during Cryo-EM Sample Blotting," *Biophys. J.*, vol. 118, no. 3, pp. 708–719, 2020, doi: 10.1016/j.bpj.2019.12.017.

[102]   E. D'Imprima, D. Floris, M. Joppe, R. Sánchez, M. Grininger, and W. Kühlbrandt, "Protein denaturation at the air-water interface and how to prevent it," *Elife*, vol. 8:e42747, 2019.

[103]   S. Li, "Detergents and alternatives in cryo-EM studies of membrane proteins," *Acta Biochim. Biophys. Sin. (Shanghai).*, vol. 54, no. 8, pp. 1049–1056, 2022, doi: 10.3724/abbs.2022088.

[104]   E. Ahn *et al.*, "Batch Production of High-Quality Graphene Grids for Cryo-EM: Cryo-EM Structure of Methylococcus capsulatus Soluble Methane Monooxygenase Hydroxylase," *ACS Nano*, vol. 17, no. 6, pp. 6011–6022, 2023, doi: 10.1021/acsnano.3c00463.

[105]   R. F. Thompson, M. G. Iadanza, E. L. Hesketh, S. Rawson, and N. A. Ranson, "Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy," *Nat. Protoc.*, vol. 14, no. 1, pp. 100–118, 2019, doi: 10.1038/s41596-018-0084-8.

[106]   I. Drulyte *et al.*, "Approaches to altering particle distributions in cryo-electron microscopy sample preparation," *Acta Crystallogr. Sect. D Struct. Biol.*, vol. 74, no. 6, pp. 560–571, 2018, doi: 10.1107/S2059798318006496.

[107]   D. Kampjut, J. Steiner, and L. A. Sazanov, "Cryo-EM grid optimization for membrane proteins," *iScience*, vol. 24, no. 3, p. 102139, 2021, doi: 10.1016/j.isci.2021.102139.

[108]   A. Merk *et al.*, "Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery," *Cell*, vol. 165, no. 7, pp. 1698–1707, 2016, doi: 10.1016/j.cell.2016.05.040.

[109]   F. Weissmann *et al.*, "BiGBac enables rapid gene assembly for the expression of large

multisubunit protein complexes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 19, pp. E2564–E2569, 2016, doi: 10.1073/pnas.1604935113.

[110] A. Sonn-Segev *et al.*, "Quantifying the heterogeneity of macromolecular machines by mass photometry," *Nat. Commun.*, vol. 11, no. 1, pp. 1–10, 2020, doi: 10.1038/s41467-020-15642-w.

[111] S. Lee *et al.*, "The structure of ClpB: A molecular chaperone that rescues proteins from an aggregated state," *Cell*, vol. 115, no. 2, pp. 229–240, 2003, doi: 10.1016/S0092-8674(03)00807-9.

[112] Z. Yu, M. D. Gonciarz, W. I. Sundquist, C. P. Hill, and G. J. Jensen, "Cryo-EM Structure of Dodecameric Vps4p and Its 2:1 Complex with Vta1p," *J. Mol. Biol.*, vol. 377, no. 2, pp. 364–377, 2008, doi: 10.1016/j.jmb.2008.01.009.

[113] S. Camier, O. Gabrielsen, R. Baker, and A. Sentenac, "A split binding site for transcription factor tau on the tRNA3Glu gene.," *EMBO J.*, vol. 4, no. 2, pp. 491–500, 1985, doi: 10.1002/j.1460-2075.1985.tb03655.x.

[114] O. Lefebvre, J. Ruth, and A. Sentenac, "A mutation in the largest subunit of yeast TFIIIC affects tRNA and 5 S RNA synthesis. Identification of two classes of suppressors," *J. Biol. Chem.*, vol. 269, no. 37, pp. 23374–23381, 1994, doi: 10.1016/s0021-9258(17)31663-0.

[115] G. Male *et al.*, "Architecture of TFIIIC and its role in RNA polymerase III pre-initiation complex assembly," *Nat. Commun.*, vol. 6, pp. 1–11, 2015, doi: 10.1038/ncomms8387.

[116] K. Van Bortle and V. G. Corces, "tDNA insulators and the emerging role of TFIIIC in genome organization," *Transcription*, vol. 3, no. 6, pp. 277–284, 2012, doi: 10.4161/trns.21579.

[117] D. Canella *et al.*, "A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver," *Genome Res.*, vol. 22, no. 4, pp. 666–680, 2012, doi: 10.1101/gr.130286.111.

[118] Z. Wang and R. G. Roeder, "TFIIIC1 Acts through a Downstream Region To Stabilize TFIIIC2 Binding to RNA Polymerase III Promoters," *Mol. Cell. Biol.*, vol. 16, no. 12, pp. 6841–6850, 1996, doi: 10.1128/mcb.16.12.6841.

[119] Y. Shen, M. Igo, P. Yalamanchili, A. J. Berk, and A. Dasgupta, "DNA binding domain and subunit interactions of transcription factor IIIC revealed by dissection with poliovirus 3C protease," *Mol. Cell. Biol.*, vol. 16, no. 8, pp. 4163–4171, 1996, doi: 10.1128/mcb.16.8.4163.

[120] D. Wu and G. Piszczek, "Standard Protocol for Mass Photometry Experiments," *Physiol. Behav.*, vol. 50, no. 3–4, pp. 403–409, 2021, doi: 10.1007/s00249-021-01513-9.Standard.

[121] R. Arrebola *et al.*, "τ91, an Essential Subunit of Yeast Transcription Factor IIIC, Cooperates with τ138 in DNA Binding," *Molecular and Cellular Biology*, vol. 18, no. 1. pp. 1–9, 1998. doi: 10.1128/mcb.18.1.1.

[122] R. Sanchez-Garcia, J. Gomez-Blanco, A. Cuervo, J. M. Carazo, C. O. S. Sorzano, and J. Vargas, "DeepEMhancer: a deep learning solution for cryo-EM volume post-processing," *Commun. Biol.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.1038/s42003-021-02399-1.

[123] R. Evans *et al.*, "Protein complex prediction with AlphaFold-Multimer," *bioRxiv*

2021.10.04.463034, 2002, doi: https://doi.org/10.1101/2021.10.04.463034.

[124] F. Van Den Ent and J. Löwe, "RF cloning: A restriction-free method for inserting target genes into plasmids," *J. Biochem. Biophys. Methods*, vol. 67, no. 1, pp. 67–74, 2006, doi: 10.1016/j.jbbm.2005.12.008.

[125] D. N. Mastronarde, "Automated electron microscope tomography using robust prediction of specimen movements," *J. Struct. Biol.*, vol. 152, no. 1, pp. 36–51, 2005, doi: 10.1016/j.jsb.2005.07.007.

[126] D. Tegunov and P. Cramer, "Real-time cryo-electron microscopy data preprocessing with Warp," *Nat. Methods*, vol. 16, no. 11, pp. 1146–1152, 2019, doi: 10.1038/s41592-019-0580-y.

[127] J. Zivanov *et al.*, "New tools for automated high-resolution cryo-EM structure determination in RELION-3," *Elife*, vol. 7, pp. 1–22, 2018, doi: 10.7554/eLife.42166.

[128] S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard, "MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy," *Nat. Methods*, vol. 14, no. 4, pp. 331–332, 2017, doi: 10.1038/nmeth.4193.

[129] K. Zhang, "Gctf: Real-time CTF determination and correction," *J. Struct. Biol.*, vol. 193, no. 1, pp. 1–12, 2016, doi: 10.1016/j.jsb.2015.11.003.

[130] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, "CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination," *Nat. Methods*, vol. 14, no. 3, pp. 290–296, 2017, doi: 10.1038/nmeth.4169.

[131] T. Bepler *et al.*, "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs," *Nat. Methods*, vol. 16, no. 11, pp. 1153–1160, 2019, doi: 10.1038/s41592-019-0575-8.

[132] E. D. Zhong, T. Bepler, B. Berger, and J. H. Davis, "CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks," *Nat. Methods*, vol. 18, no. 2, pp. 176–185, 2021, doi: 10.1038/s41592-020-01049-4.

[133] A. Casañal, B. Lohkamp, and P. Emsley, "Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data," *Protein Sci.*, vol. 29, no. 4, pp. 1069–1078, 2020, doi: 10.1002/pro.3791.

[134] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021, doi: 10.1038/s41586-021-03819-2.

[135] E. F. Pettersen *et al.*, "UCSF ChimeraX: Structure visualization for researchers, educators, and developers," *Protein Sci.*, vol. 30, no. 1, pp. 70–82, 2021, doi: 10.1002/pro.3943.

[136] R. Evans *et al.*, "Protein complex prediction with AlphaFold-Multimer," *bioRxiv Prepr.*, 2022, doi: https://doi.org/10.1101/2021.10.04.463034.

[137] P. V. Afonine *et al.*, "Real-space refinement in PHENIX for cryo-EM and crystallography," *Acta Crystallogr. Sect. D Struct. Biol.*, vol. 74, no. 6, pp. 531–544, 2018, doi: 10.1107/S2059798318006551.

[138] C. J. Williams *et al.*, "MolProbity: More and better reference data for improved all-atom structure validation," *Protein Sci.*, vol. 27, no. 1, pp. 293–315, Jan. 2018, doi: 10.1002/pro.3330.

[139] E. Krissinel and K. Henrick, "Inference of Macromolecular Assemblies from Crystalline State," *J. Mol. Biol.*, vol. 372, no. 3, pp. 774–797, 2007, doi: 10.1016/j.jmb.2007.05.022.

[140] N. M. Luscombe, R. A. Laskowski, and J. M. Thornton, "NUCPLOT: A program to generate schematic diagrams of protein-nucleic acid interactions," *Nucleic Acids Res.*, vol. 25, no. 24, pp. 4940–4945, 1997, doi: 10.1093/nar/25.24.4940.

[141] L. Holm, "Dali server: structural unification of protein families," *Nucleic Acids Res.*, vol. 50, no. W1, pp. W210–W215, 2022, doi: 10.1093/nar/gkac387.

[142] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, and K. Zakrzewska, "Conformational analysis of nucleic acids revisited: Curves+," *Nucleic Acids Res.*, vol. 37, no. 17, pp. 5917–5929, Sep. 2009, doi: 10.1093/nar/gkp608.

[143] D. Kimanius, L. Dong, G. Sharov, T. Nakane, and S. H. W. Scheres, "New tools for automated cryo-EM single-particle analysis in RELION-4.0," *Biochem. J.*, vol. 478, no. 24, pp. 4169–4185, 2021, doi: 10.1042/BCJ20210708.