Transcription in eukaryotes

Comparison of eukaryotic and prokaryotic RNA polymerases

Eukaryotes: **Three polymerase** transcribes different class of genes:

Pol I-major rRNA (18S, 5.8S, 28S) genes

Pol II-mRNA genes

Pol III-tRNA, minor (5S) rRNA & snRNA genes (U6)

Prokaryotes: one polymerase transcribes all genes

Table 10.1 Roles of Eukaryotic RNA Polymerases

RNA Polymerase	Cellular RNAs Synthesized	Mature RNA (Vertebrate)
I	Large rRNA precursor	28S, 18S, and 5.8S rRNAs
II	hnRNAs	mRNAs
	snRNAs	snRNAs
III	5S rRNA precursor	5S rRNA
	tRNA precursors	tRNAs
	U6 snRNA (precursor?)	U6 snRNA
	7SL RNA (precursor?)	7SL RNA
	7SK RNA (precursor?)	7SK RNA

Table 10.2 Human and Yeast RNA Polymerase II Subunits

Subunit	Yeast Gene	Yeast Protein (kD)	Features
hRPB1	*RPB1*	192	Contains CTD; binds DNA; involved in start site selection; β′ ortholog
hRPB2	*RPB2*	139	Contains active site; involved in start site selection, elongation rate; β ortholog
hRPB3	*RPB3*	35	May function with Rpb11 as ortholog of the α dimer of prokaryotic RNA polymerase
hRPB4	*RPB4*	25	Subcomplex with Rpb7; involved in stress response
hRPB5	*RPB5*	25	Shared with Pol I, II, III; target for transcriptional activators
hRPB6	*RPB6*	18	Shared with Pol I, II, III; functions assembly and stability
hRPB7	*RPB7*	19	Forms subcomplex with Rpb4 that preferentially binds during stationary phase
hRPB8	*RPB8*	17	Shared with Pol I, II, III; has oligonucleotide/oligosaccharide-binding domain
hRPB9	*RPB9*	14	Contains zinc ribbon motif that may be involved in elongation: functions in start site selection
hRPB10	*RPB10*	8	Shared with Pol I, II, III
hRPB11	*RPB11*	14	May function with Rpb3 as ortholog of the α dimer of prokaryotic RNA polymerase
hRPB12	*RPB12*	8	Shared with Pol I, II, III

Polymerase II

10 subunits are placed in 3 groups:

- **Core** – (3 of the subunits) - related in structure and function to bacterial core subunits
- **Common** – (5 of the subunits) - found in all 3 nuclear RNA polymerases in yeast
- **Nonessential subunits** – (2 of the subunits) - conditionally dispensable for enzymatic activity

Core Subunits

- Three polypeptides - Rpb1, Rpb2, Rpb3 -absolutely required for enzyme activity
- These are homologous to β'-, β-, and α-subunits
- Both **Rpb1** and β'-subunit binds DNA
- **Rpb2** and β-subunit are at or near the nucleotide-joining active site
- **Rpb3** does not resemble α-subunit
  - There is one 20-amino acid subunit of great similarity
  - 2 subunits are about same size - same stoichiometry

Common Subunits

- There are five common subunits
  - Rpb5
  - Rpb6
  - Rpb8
  - Rpb10
  - Rpb12
- Little known about function
- They are all found in all 3 polymerases
- Suggests play roles fundamental in transcription

Subunits Nonessential for Elongation

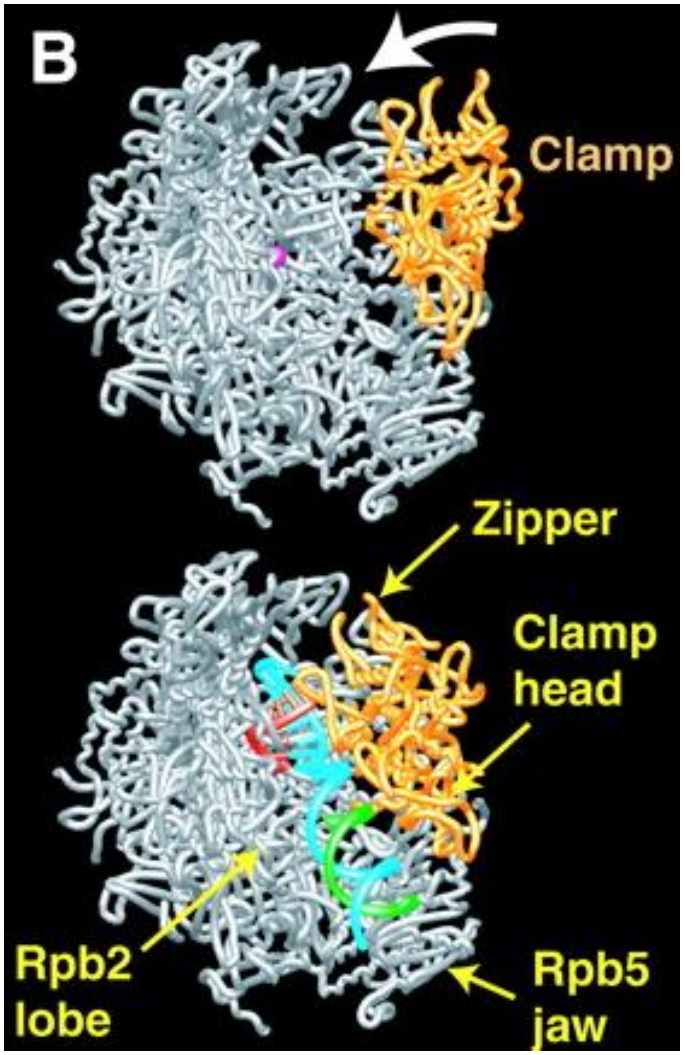
- Rpb4 and Rpb7
  - **Dissociate fairly easily** from polymerase
  - Might **shuttle from** one polymerase II to another
  - **Rpb4 may help anchor Rpb7** to the enzyme
  - Mutants **without Rpb4 and Rpb7** transcribes well- but cannot initiate at a real promoter
- Rpb7 is an essential subunit for initiation

The Three-Dimensional Structure of RNA Polymerase II

- Structure of eukaryotic polymerase II - have **deep cleft** that accepts a **linear DNA template** from one end to another
- Catalytic center lies at the bottom of the cleft and contains a **Mg^{2+} ion**
- **Upper jaw** – Rpb1+Rpb9 and **lower jaw** – Rpb5
- Geometry allows enough space for:
  - TFIID to bind at the TATA box of the promoter
  - TFIIB to link the polymerase to TFIID
  - Places polymerase correctly to initiate transcription

Differences between non-transcribing and transcribing structures of RNA Pol II

- The non-transcribing complex is more “open”, which faciliates DNA binding
- Upon the binding to promoter, the structure, called “clamp” closes the complex like a lid



Comparison of eukaryotic and prokaryotic promoter recognition

Eukaryotes: General transcription factors (GTFs).

TFI factors for RNAP I

TFII factors for RNAP II

TFIII factors for RNAP III

Prokaryotes: σ factors

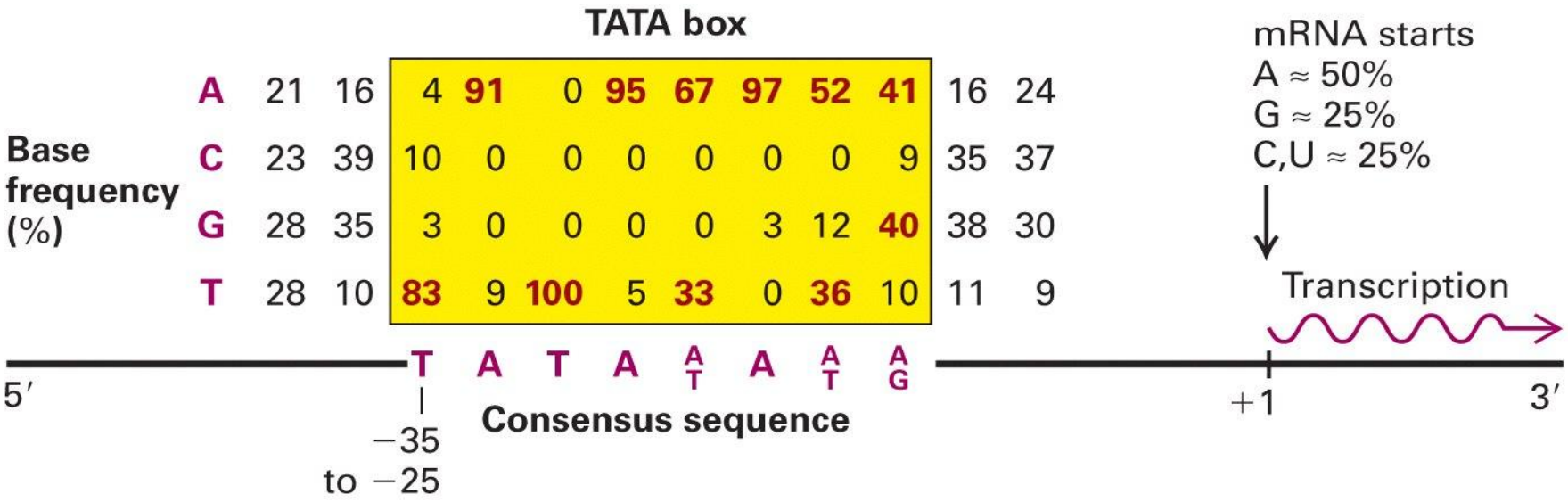
Eukaryotic promoters

RNA polymerase II core promoters are made up of combinations of 4 different sequence elements

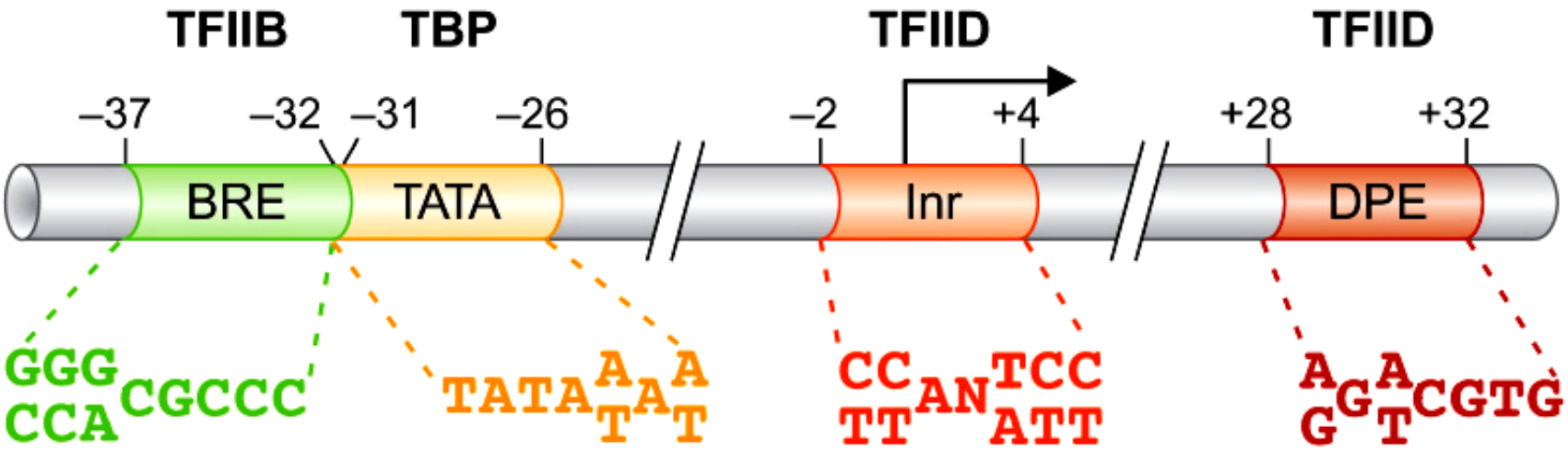
Eukaryotic core promoter (~40 nt):

The minimal set of sequence elements required for accurate transcription initiation by the RNA Pol II machinery in vitro

The TATA box


TATA box
Base frequency (%)
A 21 16 4 91 0 95 67 97 52 41 16 24
C 23 39 10 0 0 0 0 0 0 9 35 37
G 28 35 3 0 0 0 0 3 12 40 38 30
T 28 10 83 9 100 5 33 0 36 10 11 9
5′
T A T A A/T A A/T A/G
3′
−35 to −25
Consensus sequence
+1
mRNA starts
A ≈ 50%
G ≈ 25%
C,U ≈ 25%
Transcription

RNA Pol II core promoter

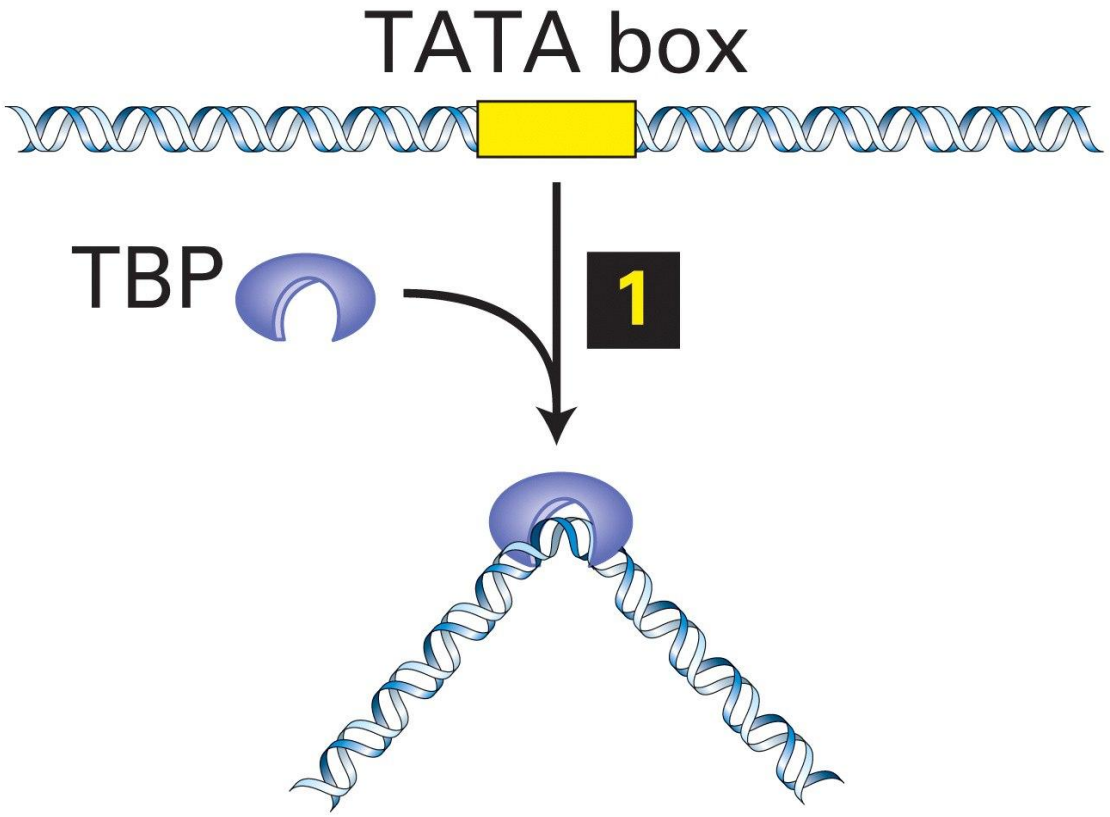


- TFIIB recognition element (BRE)
- The TATA element/box
- Initiator (Inr)
- The downstream promoter element (DPE)

General transcription factors

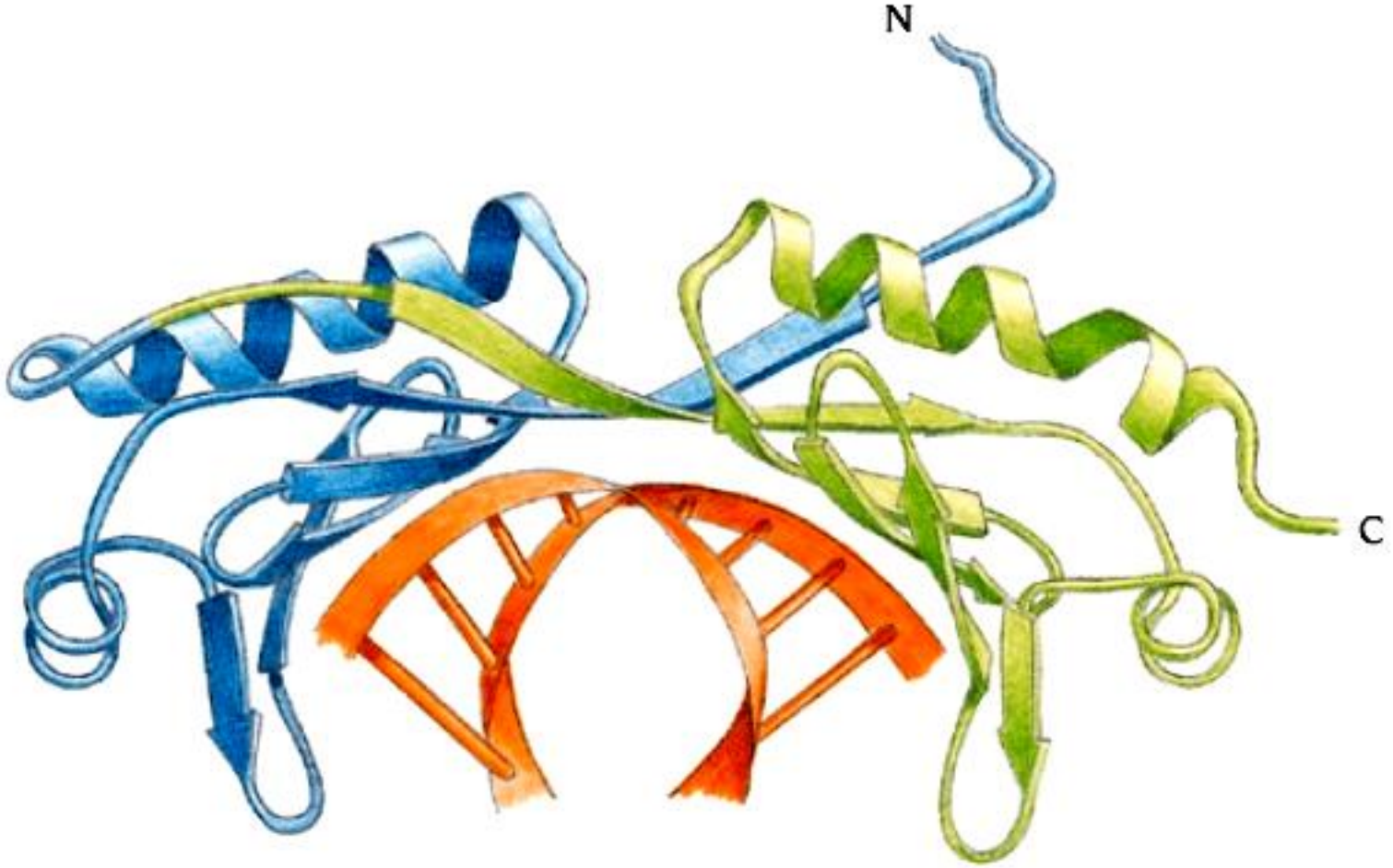
- **G**eneral **t**ranscription **f**actors are required for transcription in eukaryotes from all genes
- GTFs assist RNA Pol II in transcription initiation
- GTFs are designated TFIIA, TFIIB,... and most of them are multimeric proteins
- Equivalent GTFs are highly conserved among the eukaryotes
- In prokaryotes, only one general transcription factor, known as σ factor is required

TFIID

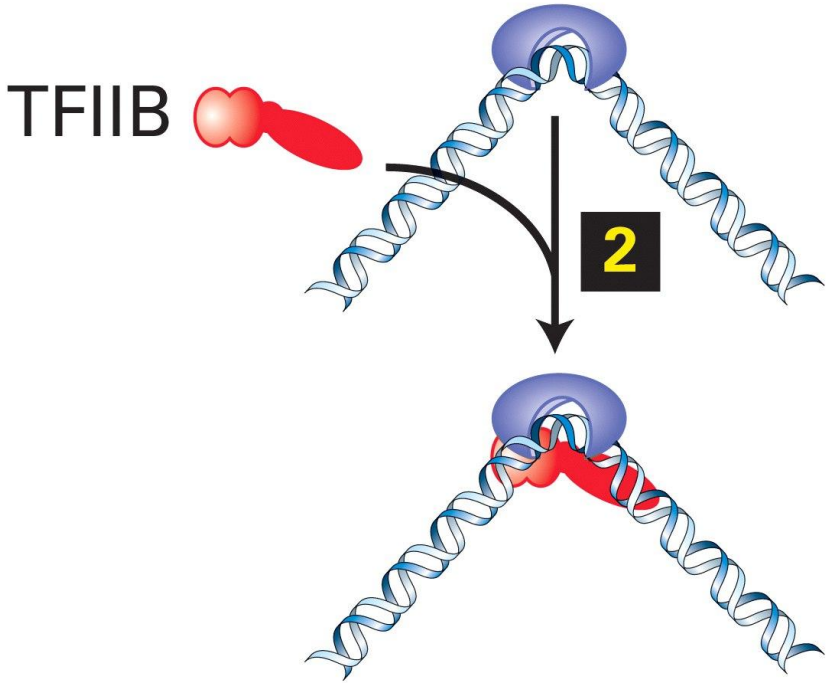


- TFIID is composed of 14 subunits, one of them being **T**ATA box **b**inding **p**rotein, TBP
- Functions: promoter recognition, TFIIB recruitment

The TATA box binding protein (TBP)



TFIIB

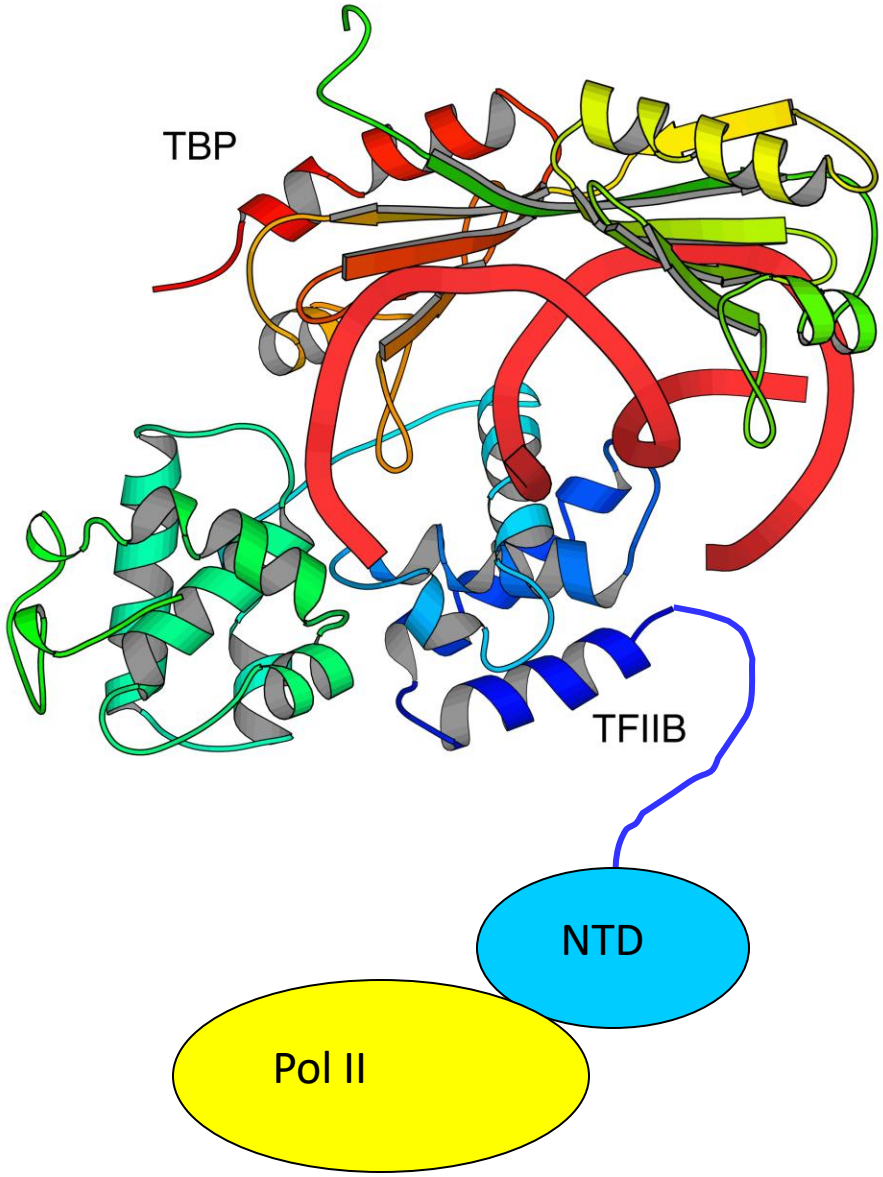


Functions:

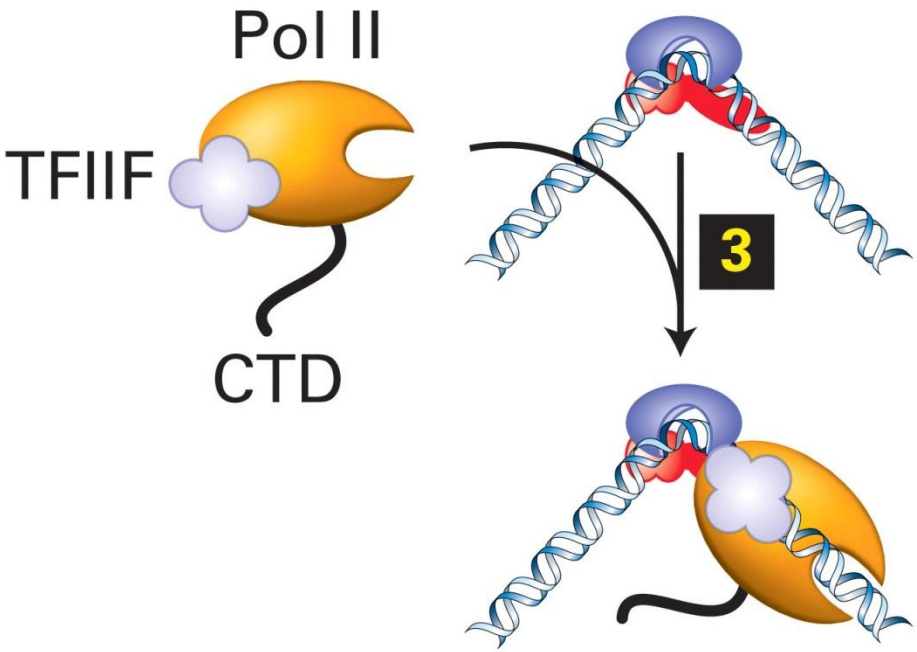
- Start site selection for Pol II
- TFIIF-RNA PolII complex recruitment
- Some TFIIB mutants result in a shift of transcription start site
- Under certain conditions (BRE element promoter) Pol II together with TFIID and TFIIB can form the minimal initiation complex. At most promoters, however, TFIIE, F and H are necessary

C-terminal domain (CTD) of TFIIB

- CTD of TFIIB interacts with both TBP and DNA around the promoter – especially BRE element
- Rough positioning of Pol II is due to interaction of TFIIB CTD with TBP
- Fine positioning is due to interaction with DNA
- N-terminal domain of TFIIB interacts with RNA Pol II



TFIIF

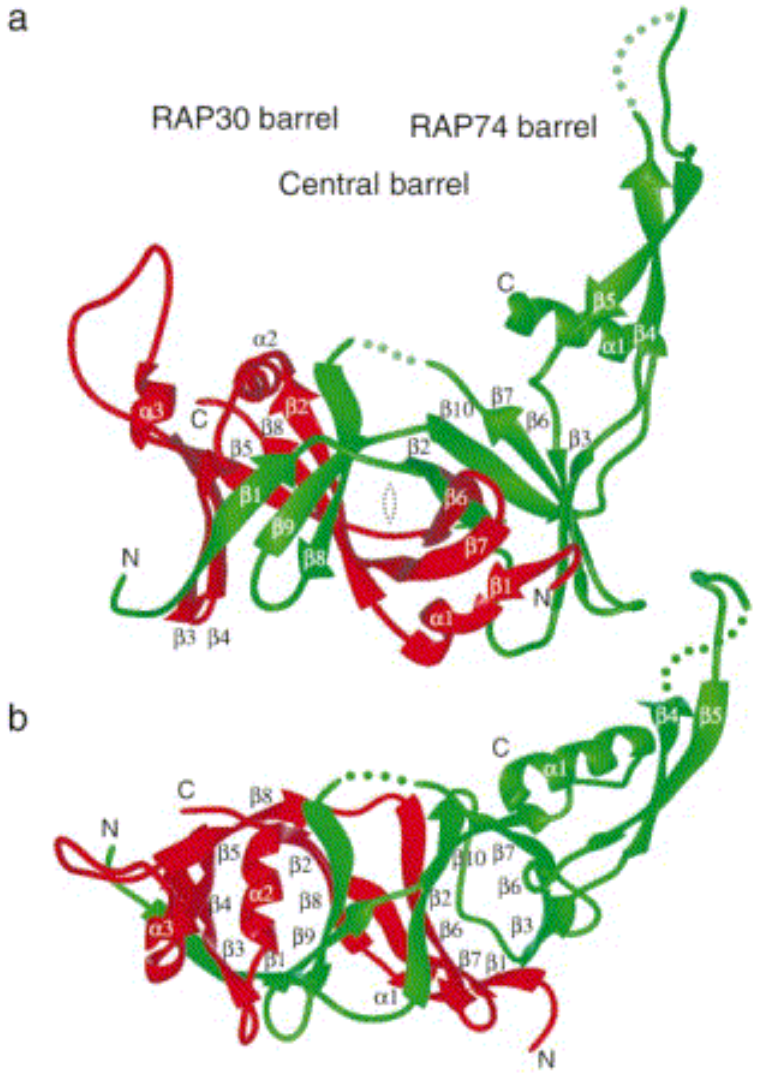


Functions:

- Recruitment of Pol II to the existing DNA-TFIID-B complex,
- Positioning Pol II over the start site
- Binding to the non-template DNA strand.
- TFIIF also reduces non-specific binding of RNA pol II to DNA.

Structure of TFIIF

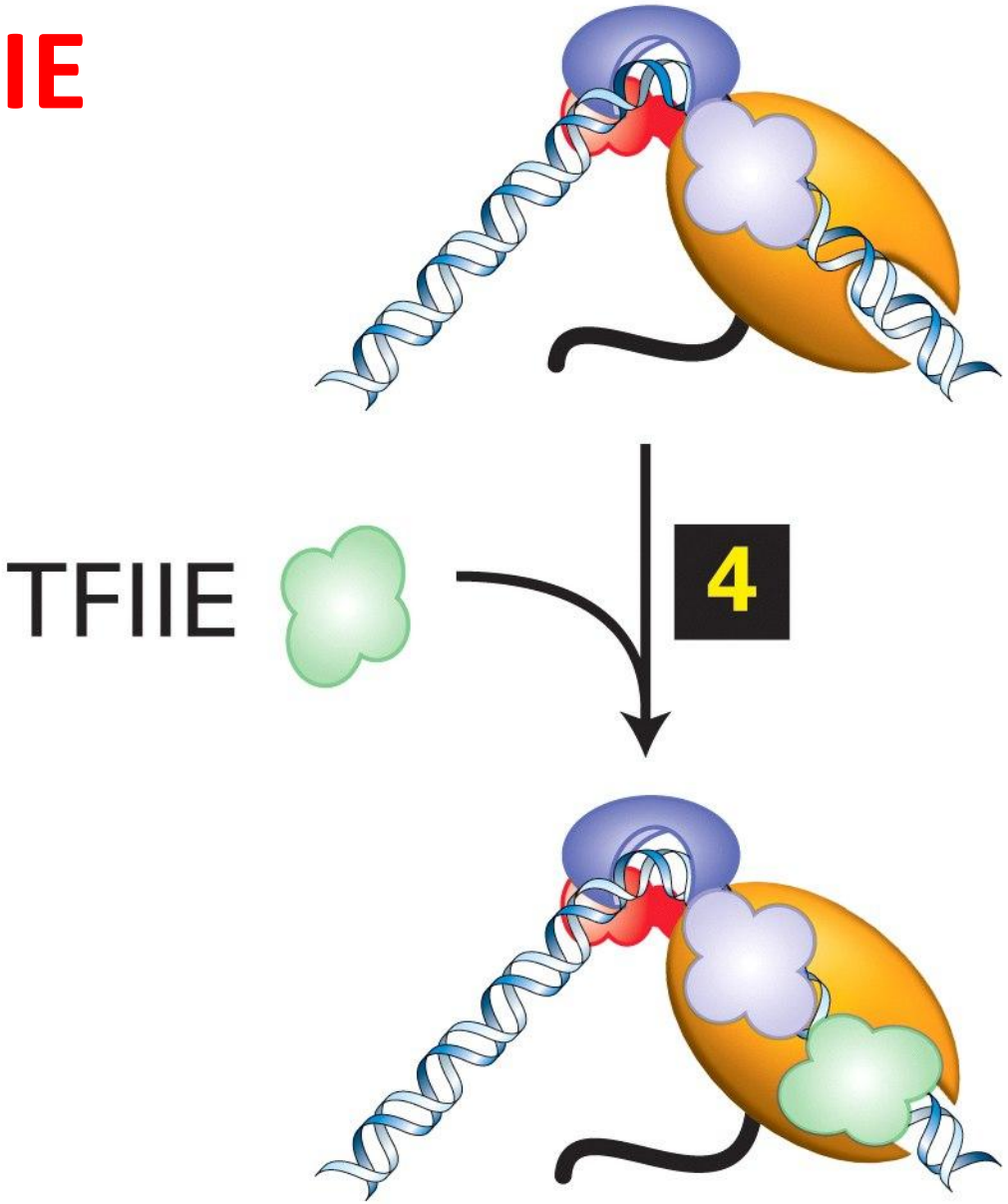
- TFIIF is a heterotetramer, made from two subunits of RAP30 and two subunits of RAP74 proteins
- RAP30-RAP74 dimer within the complex structure has an unusual triple-barrel fold



TFIIE

- TFIIE is a heterotetrameric protein ($\alpha_2\beta_2$)
- **Functions:**

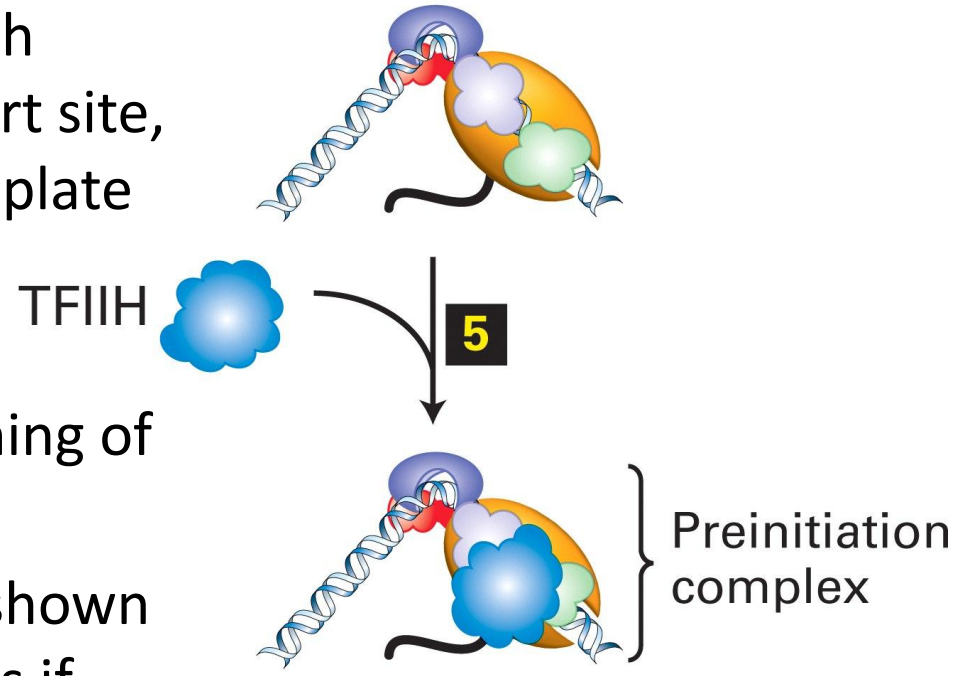
- TFIIE appears to create the docking site for next transcription factor, TFIIH.
- TFIIE also modulates TFIIH enzymatic activities
- In addition TFIIE enhances promoter melting.

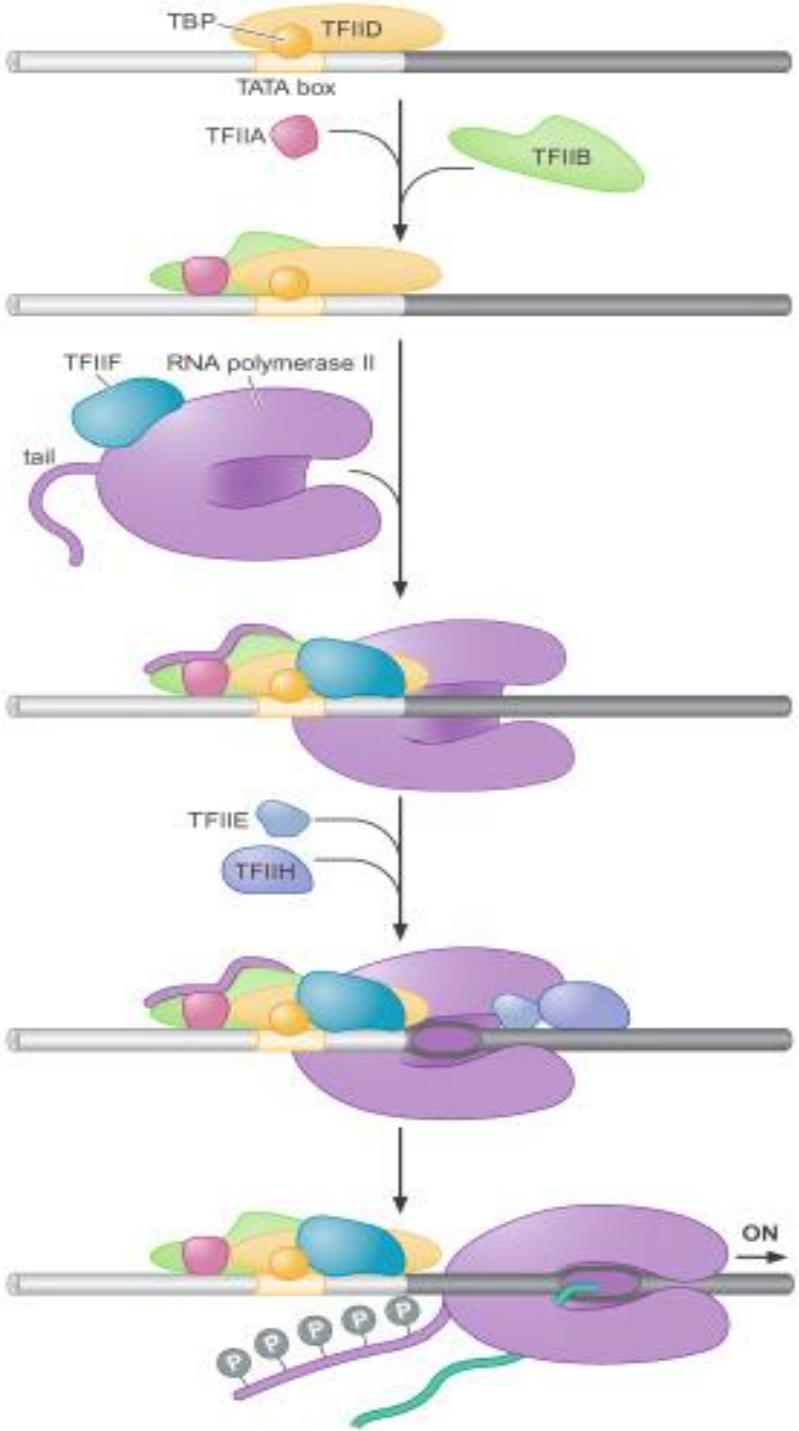


TFIIH

- TFIIH is a multimeric protein, composed of 9 subunits, some of them with distinct enzymatic activities
- **Functions:**

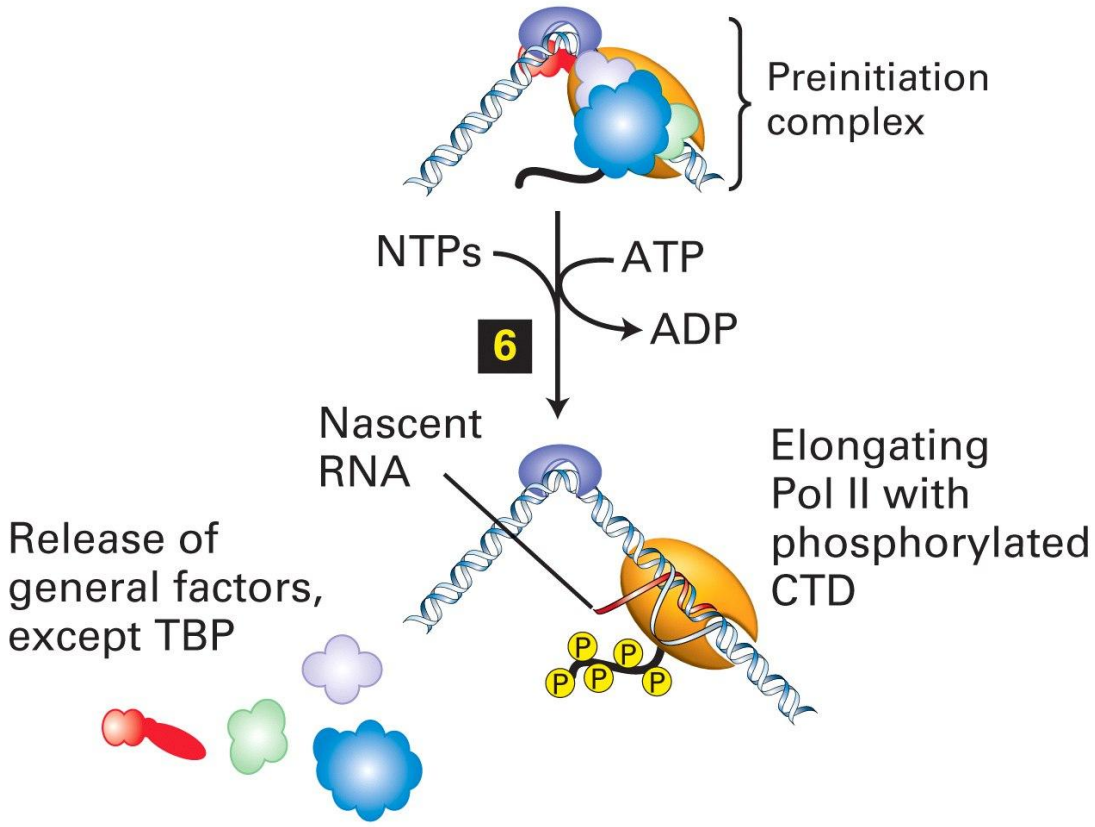
- TFIIH has a helicase activity, which unwinds the DNA duplex at a start site, allowing Pol II to bind to the template strand.
- TFIIH also has a kinase ativity, it phosphorylates PolII in the begining of elongation
- Other TFIIH subunits have been shown to recruit DNA-repairing enzymes if polymerase reaches damaged region in DNA and gets stalled





1. <u>TBP in TFIID</u> binds to the TATA box
2. <u>TFIIA and TFIIB</u> are recruited with TFIIB binding to the BRE
3. <u>RNA Pol II-TFIIF</u> complex is then recruited
4. <u>TFIIE and TFIIH</u> then bind upstream of Pol II to form the pre-initiation complex
5. Promoter melting using energy from ATP hydrolysis by TFIIH)
6. Promoter escapes after the phosphorylation of the CTD tail

Early events in elongation



- As Pol II transcribes away from the start site subunit of TFIIH phosphorylates the Pol II CTD, which results in promoter escape.
- General factors get released

TFIIA

- For transcription *in vivo*, another factor TFIIA is required
- The function of TFIIA is somewhat unclear, but it might help the other factors to bind.
- TFIIA has also shown to have some anti-repressor functions
- TFIIA is not required for transcription *in vitro.*

TAFs (TBP associated Factors)

- Apart from TBP, TFIID has 13 TAF (**T**BP **a**ssociated **f**actors) subunits
- Some of them seem to be necessary for transcription initiation from promoters, lacking the TATA box
- Other TAFs have been shown to be tissue-specific coactivators
- TAF subunits also interact with other GTFs therefore stabilyzing the complex.

Transcription elongation factors

Large number of proteins and protein complexes exist that regulate elongation of transcription (elongation factors)

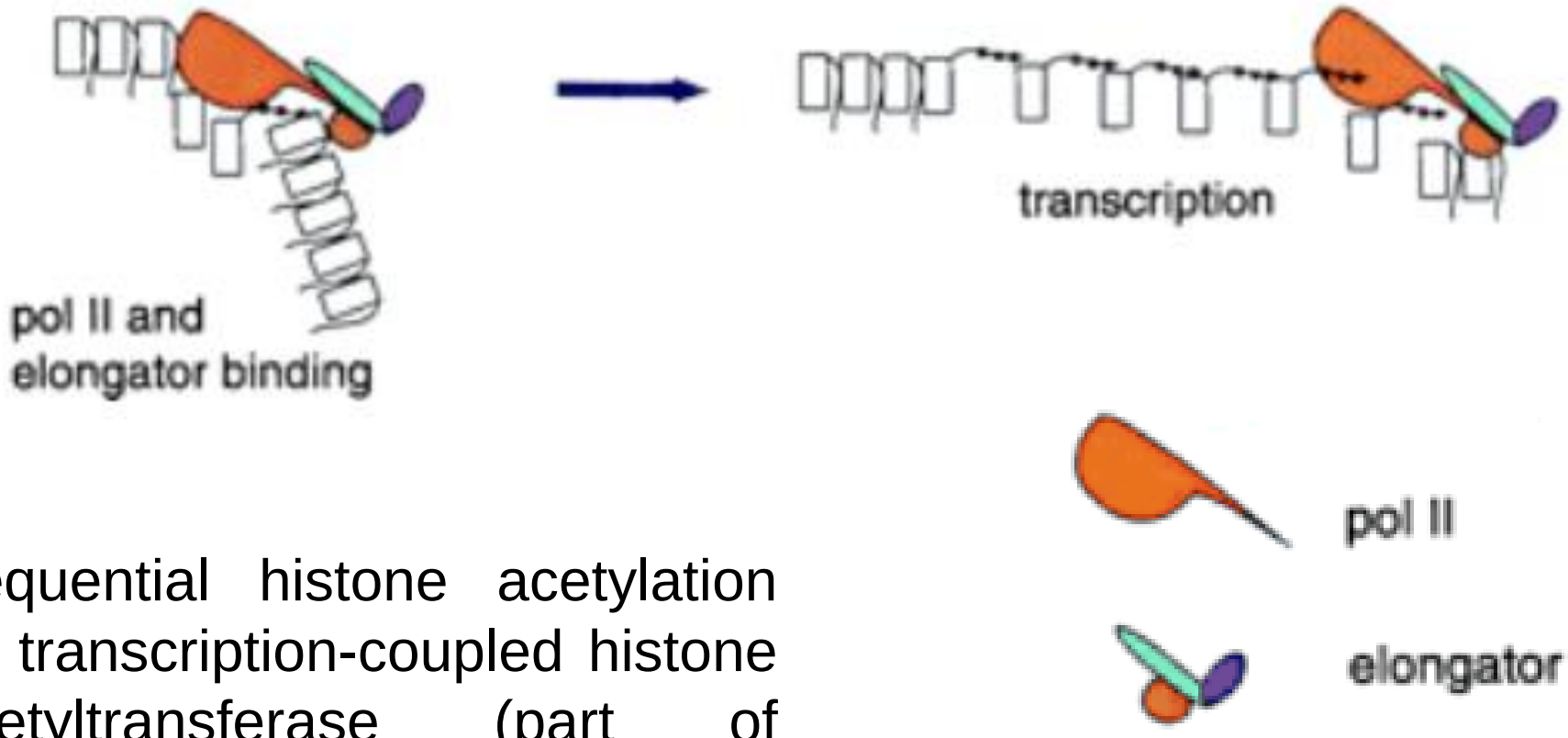
Elongation factors display following activities:

- Enable elongation through chromatin
- Suppress pausing
- Overcome transcription arrest

Some elongation factors

- **P-TEFb (**Positive Transcription Elongation Factor**)**:
  - phosphorylates CTD
  - Activates hSPT5
  - Activates TAT-SF1
- **TFIIS**:
  - Stimulates the overall rate of elongation by resolving the polymerase pausing
  - Proofreading

Elongator decondenses the chromatin



Sequential histone acetylation by transcription-coupled histone acetyltransferase (part of elongator complex)

Other elongation factors

- **ELL family** of proteins also stimulates transcription elongation through a poorly understood mechanism. There is some evidence, that ELLs target RNA polymerase for degradation upon reaching damaged regions in DNA. Degradation of PolII is followed by recruitment of DNA repairing enzymes.

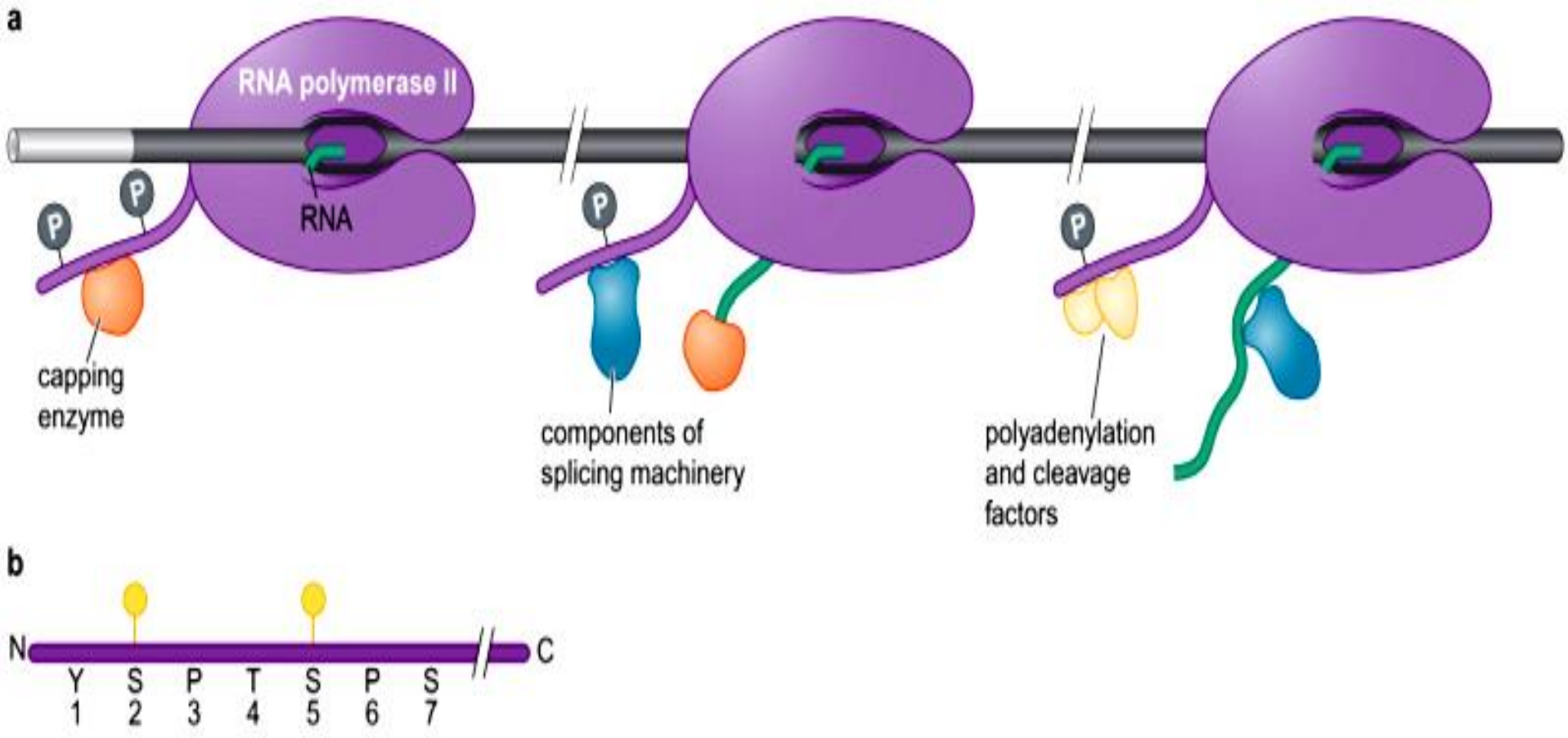
- **Negative factors NELF and DSIF** supress elongation by direct interaction with Pol II

Transition from the initiation to elongation involves the Pol II enzyme shedding most of its initiation factors (GTF and mediators) and recruiting other factors:

(1) Elongation factors: factors that stimulate elongation, such as TFIIS and hSPT5.

(2) RNA processing factors

Recruited to the C-terminal tail of the CTD of RNAP II to phosphorylate the tail for elongation stimulation, proofreading, and RNA processing like splicing and polyadenylation.



RNA processing enzymes are recruited by the tail of polymerase

RNA Polymerase is associated with a new set of protein factors required for various types of RNA processing

RNA processing:

- Capping of the 5’ end of the RNA
- Splicing of the introns (most complicated)
- Poly adenylation of the 3’ end

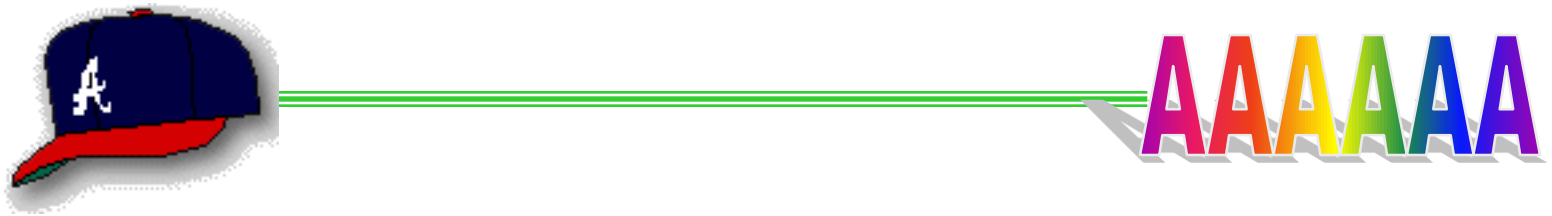
Elongation, termination of transcription, and RNA processing are interconnected/ coupled to ensure the coordination of these events

Function of 5´cap

- Protection from degradation
- Increased translational efficiency
- Transport to cytoplasm
- Splicing of first intron

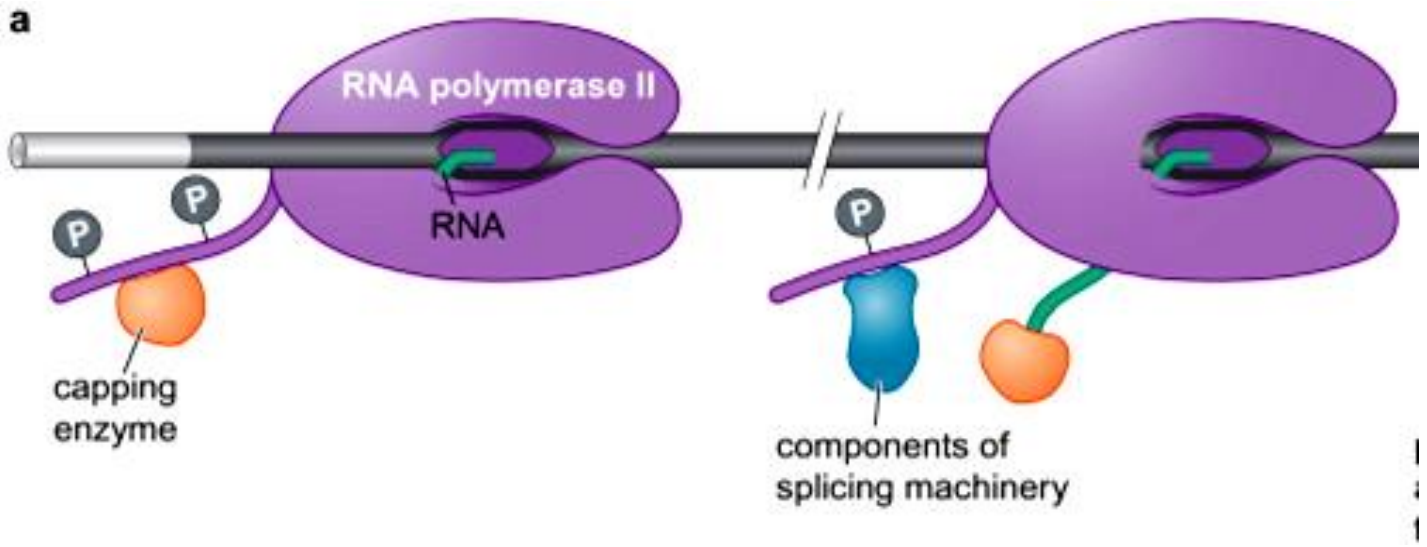
Function of poly(A) tail

- Increased mRNA stability
- Increased translational efficiency
- Splicing of last intron



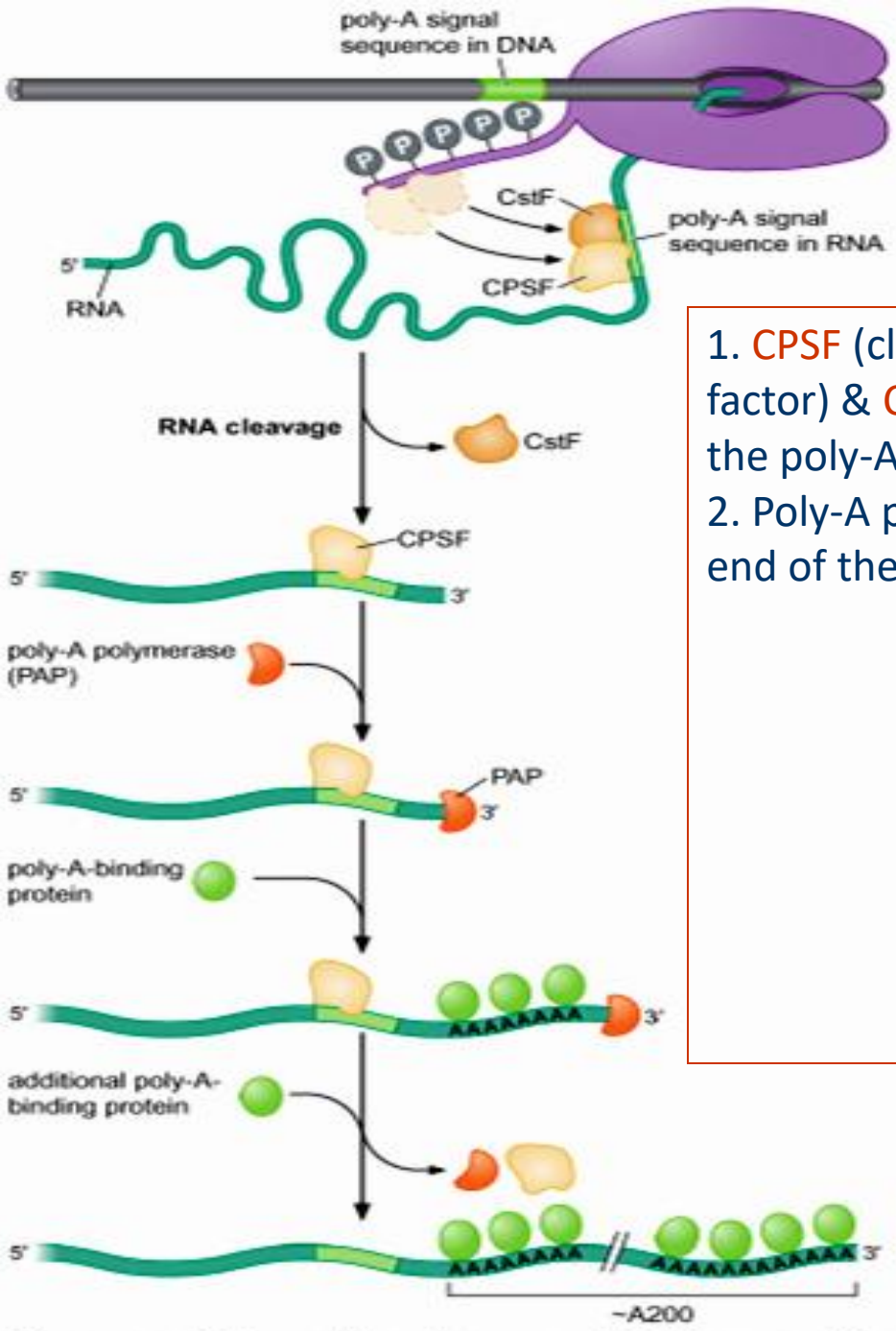
Splicing: Joining the protein coding sequences

- Dephosphorylation of Ser5 within the CTD tail leads to dissociation of capping machinery
- Further phosphorylation of Ser2 recruits the splicing machinery



3’ end polyadenylation

- Linked with the termination of transcription
- The CTD tail is involved in recruiting the polyadenylation enzymes
- The transcribed poly-A signal triggers the reactions
  1. Cleavage of the message
  2. Addition of poly-A
  3. Termination of transcription



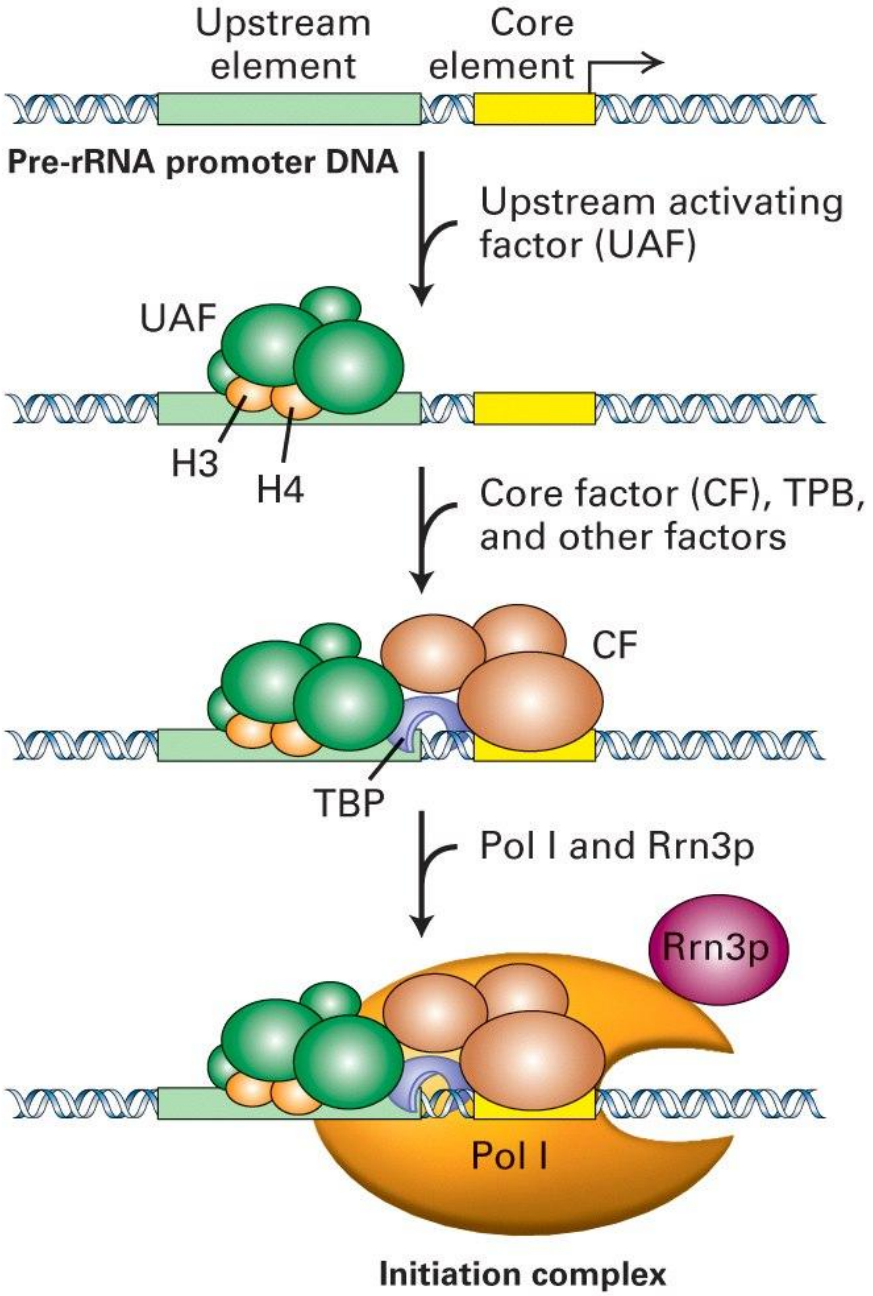
1. CPSF (cleavage and polyadenylation specificity factor) & CstF (cleavage stimulation factor) bind to the poly-A signal, leading to the RNA cleavage
2. Poly-A polymerase (PAP) adds ~ 200 As at the 3' end of the RNA, using ATP as a substrate

Polyadenylation and termination

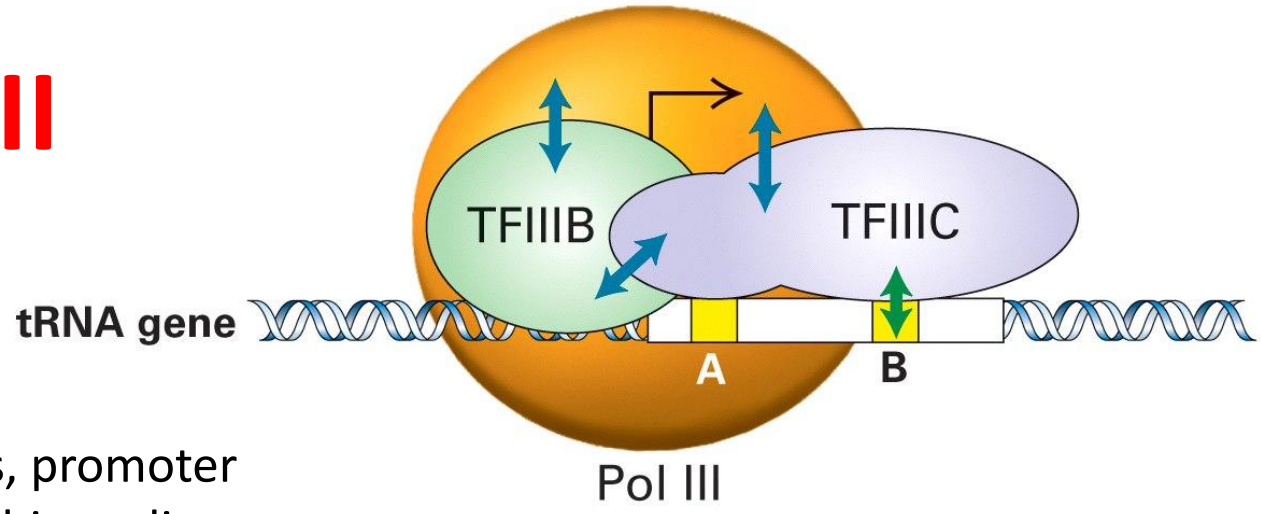
Other eukaryotic transcription systems

- RNA pol I (pre-rRNA)
- RNA pol III (tRNA, 5S rRNA)
- Mitochondrial RNA polymerase
- Chloroplast RNA polymerase

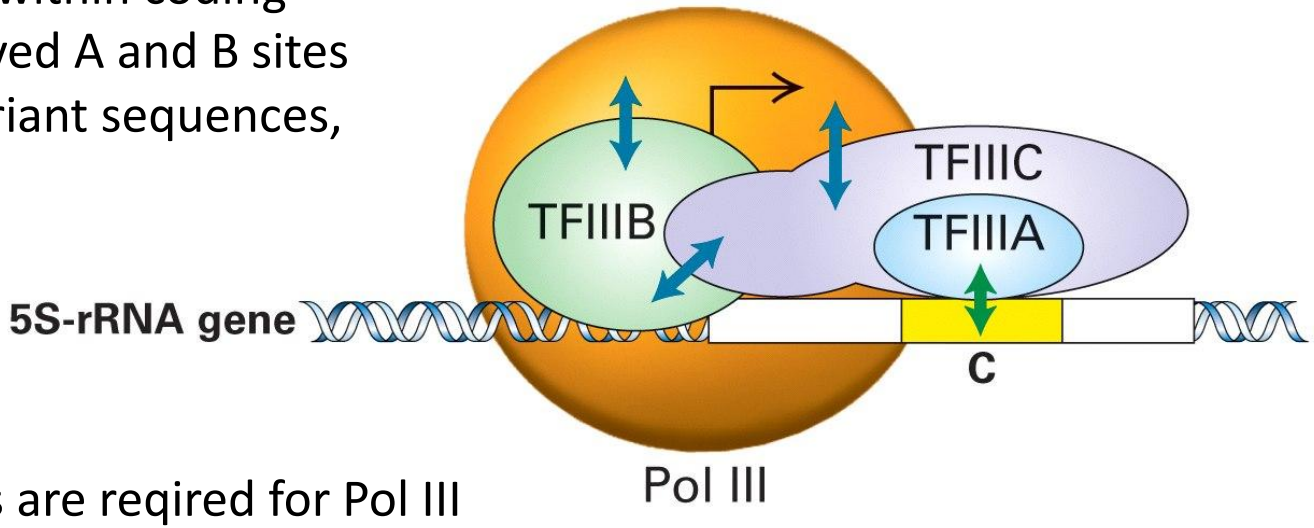
RNA Pol I


Upstream element
Core element
Pre-rRNA promoter DNA
Upstream activating factor (UAF)
UAF
H3
H4
Core factor (CF), TPB, and other factors
CF
TBP
Pol I and Rrn3p
Rrn3p
Pol I
Initiation complex

RNA pol III



Unlike other polymerases, promoter sequences lie entirely within coding sequence. The conserved A and B sites also code for two invariant sequences, observed in all tRNAs



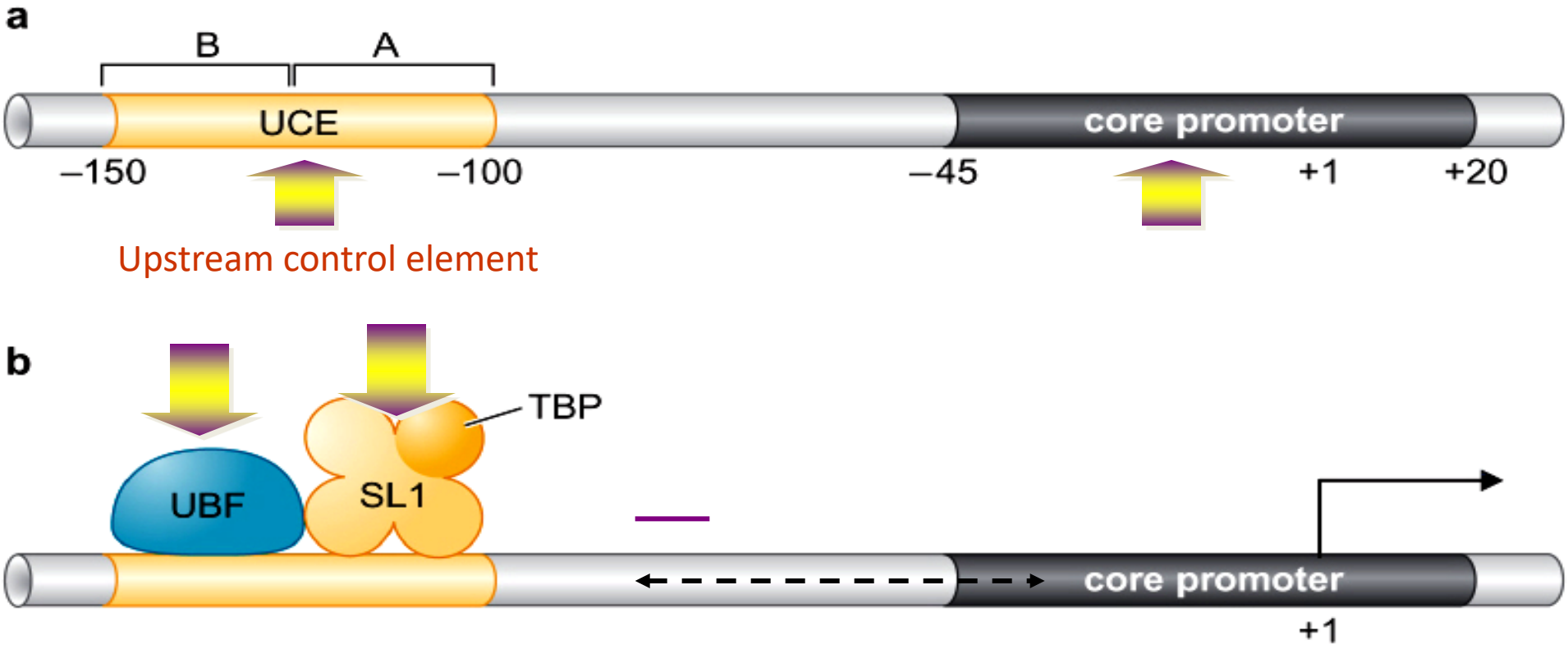
Three initiation factors are reqired for Pol III initiation. TFIIIA is required only for 5S rRNA transcription. TFIIIB has three subunits, one of them similar to TFIIB and another one being TBP (same as for Pol I and II)

RNA Pol I & III recognize distinct promoters , using distinct sets of transcription factors, but still require TBP

- Pol I: transcribes rRNA precursor encoding gene (multi-copy gene)

- Pol III: transcribes tRNA genes, snRNA genes and 5S rRNA genes

Pol I promoter recognition

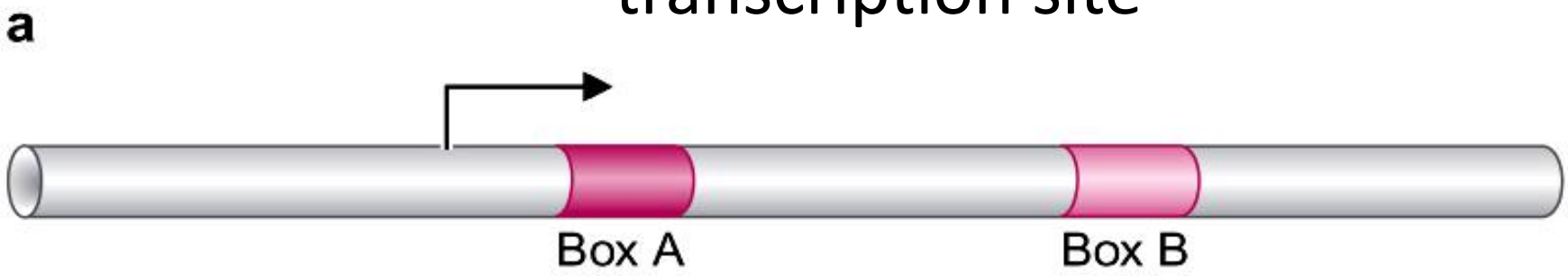


UBF binds to the upstream of UCE, bring SL1 to the downstream part of UCE. SL1 in turn recruits RNAP I to the core promoter for transcription

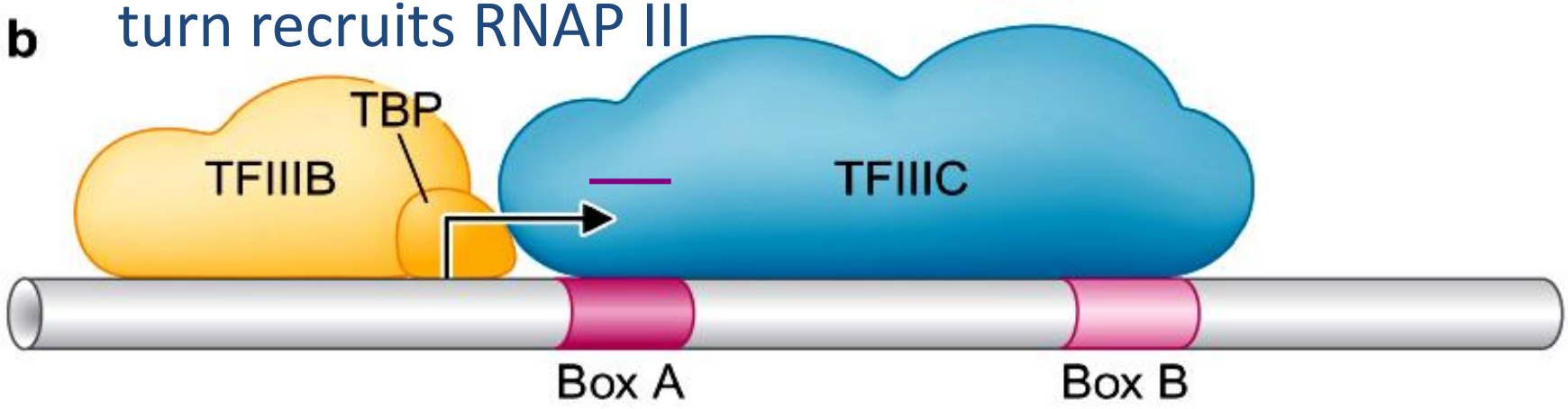
Pol I promoter region

Pol III promoter recognition

1. Different forms, 2. locates downstream of the transcription site



TFIIIC binds to the promoter, recruiting TFIIIB, which in turn recruits RNAP III



Pol III core promoter