RNA polymerase II
The central enzyme of gene expression
Enzymatic function

- **Enzymatic reaction:** NTP $\rightarrow$ RNA + PPi (1969)
  - $\text{RNA}_n + \text{NTP} + (\text{Mg}^{++} + \text{templat}) = \text{RNA}_{n+1} + \text{PP}_i$
  - Processive - can transcribe $10^6$ bp template without dissociation
  - mRNA levels can vary with a factor of $10^4$

- **Central role:** unwind the DNA double helix, polymerize RNA, and proofread the transcript

- **RNAPII** assembles into larger initiation and elongation complexes, capable of promoter recognition and response to regulatory signals
Polymerization reaction

1. Initiation
   - PIC assembly (pre-initiation complex)
   - Open complex formation
   - Promoter clearance

2. Elongation - transition to stable TEC
   - (transcription elongation complex)

3. Termination
Subunit structure

- **Composition and stoichiometry**
  - 12 polypeptides
  - 2 large (220 and 150 kDa) + 10 small (10 - 45 kDa)
  - Yeast: 10 essential, 2 non-essential
  - Phosphorylated subunits: RPB1 and RPB 6

- **Highly conserved between eukaryotes**
  - Several subunits in yeast RNAPII can be functionally exchanged with mammalian subunits
Subunits of RNA polymerase II

- The yeast model

<table>
<thead>
<tr>
<th>Factor</th>
<th>Gene</th>
<th>Subunit</th>
<th>Essential?</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA polymerase II</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPB1</td>
<td>192</td>
<td>Y</td>
<td></td>
<td>Heptapeptide repeat</td>
</tr>
<tr>
<td>RPB2</td>
<td>139</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPB3</td>
<td>35</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPB4</td>
<td>25</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPB5</td>
<td>25</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPB6</td>
<td>18</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB7</td>
<td>19</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB8</td>
<td>17</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB9</td>
<td>14</td>
<td>N</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB10</td>
<td>8</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB11</td>
<td>14</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
<tr>
<td>RPB12</td>
<td>8</td>
<td>Y</td>
<td></td>
<td>Shared with PolI, II, III</td>
</tr>
</tbody>
</table>
Evolutionary conservation of Subunits of RNA polymerase II

- Core-enzyme with the active site
  - RPB1 (β´-like) binds DNA
  - RPB2 (β-like) binds NTP
  - RPB3 and RPB11 (α-like) assembly factors

- Evolutionary conserved mechanism of RNA synthesis

- Common subunits
  - RPB5, 6, 8, 10 and 12 common to RNAPI, II and III
  - Common functions?

- Unlike prokaryotic RNAP, the eukaryotic RNAPII is unable by itself to recognize promoter sequences
3D structure of RNAPII
Yeast RNAPII

- The two largest subunits, Rpb1 and Rpb2, form masses with a deep cleft between them.
- The small subunits are arranged around.
Transcription in focus

The Nobel Prize in Chemistry 2006
"for his studies of the molecular basis of eukaryotic transcription"

Roger D. Kornberg
USA
Stanford University
Stanford, CA, USA

RNA polymerase II

Transcription factors

Co-activators
Simplified structure
Simplified structure
Several important subdomains

- Channel for DNA template (downstream)
- Jaws
- Clamp
- Wall
- Active site
- Pore for NTP entry
- Channel for RNA exit
- Hybrid melting
  - fork loop 1 + rudder + lid
- Dock
- CTD
**Channel for DNA template:**
25Å channel through the enzyme

![Diagram showing the RNA groove and the enzymes strand moving through it](image)

**yRNAPII**
Jaws

- A pair of jaws that appear to grip DNA downstream of the active center.
  - Rpb5 and regions of Rpb1 and Rpb9 forms "jaws" that appear to grip the DNA
  - Both the upper and lower jaw may be mobile, opening and closing on the DNA
A clamp retains DNA

- A clamp on the DNA nearer the active center may be locked in the closed position by RNA $\rightarrow$ great stability of complexes.
  - The "clamp" = N-terminal regions of Rpb1 and Rpb6, and the C-terminal regions of Rpb2
  - This binding site is important for the great stability of a transcribing complex and processivity of transcription
A clamp retains DNA
Moving through the compartments

- DNA enters RNAPII in the first chamber (jaw-lobe module).
  - This module binds 15–20 bp of the downstream DNA without melting it.
Moving through the 2. compartment

- The DNA melts as it enters the second chamber
  - A 27-40 Å cleft that contains the active site near the point of DNA melting.
  - The first 8–9 nt of product RNA form a heteroduplex with the template DNA (hybrid).
  - At the upstream end, a wall of protein blocks extension of the RNA:DNA hybrid.
The active site

- Reaction catalyzed

- Two NTP sites: A + E
  - Addition site
  - Entry site

A funnel for substrate entry

- A pore in the protein complex above the active center may allow entry of substrates for polymerization.
The wall and the DNA-RNA hybrid site

- Transcribing polymerases have a DNA-RNA hybrid of 8-9 bp in an unwound region of DNA, with the growing end of RNA at the active site.
- The DNA-RNA hybrid can’t get longer because of an element from Rpb2 that is blocking the path.
- Because of this “wall”, the DNA-RNA hybrid must be tilted relative to the axis of the downstream DNA.
- At the upstream end of the DNA-RNA hybrid, the strands must separate.
RNA-DNA hybrid - 90°

- The DNA is unwound, with 9 bp of DNA–RNA hybrid in the active center region.
- The axis of the hybrid helix is at nearly 90° to that of the entering DNA duplex, due to the wall.

Melting the RNA-DNA hybrid

- Melting of the DNA–RNA hybrid due to the intervention of three protein loops:
  - Rudder ("ror") contacting DNA, and
  - Lid - contacting RNA. A Phe side chain serves as a wedge to maintain separation of the strands.
  - Fork loop 1 contacts base pairs 6 and 7, limiting the strand separation.

- The three loops form a strand-loop network, whose stability must drive the melting process.

RNA exit

- Groove in the RNAPII structure for RNA exit.
- Length and localization of the groove are appropriate for binding a region of RNA 10-20 nt from the active site.
- RNA in the groove at the base of the clamp could explain the great stability of transcribing complexes.
Dock

- Contact region for specific interacting GTFs
- More next lecture
Rbp7/4 - recently determined

Rbp7 acts as a wedge to lock the clamp in the closed conformation
Opening and closing of RNAP II during the transcription cycle

- **Open RNAP** during formation of PIC
  - moderate stability
  - Strand separation and placement of template in active site, transcription bubble
  - "Abortive initiation" (RNA up to 10 nt) without structural change

- **RNAP closes** during *promoter clearance* and transition to TEC
  - contacts to PIC are disrupted and new contacts with elongation factors formed
  - CTD is phosphorylated (more later)
  - Conformational change to a ternary complex of high stability
  - Closed channel around the DNA-RNA hybrid in the active site

- **RNAP opens and becomes destabilised during termination**
  - Reversal of the structural changes - opening and destabilization
CTD - C-terminal domain

- **Conserved tail on the largest subunit**: (YSPTSPS)$n$
  - Yeast $n = 26$, humans $n = 52$
  - Hydrophilic exposed tail

- **Unique for RNAPII**

- **Essential function in vivo**
  - $\Delta >50\%$ lethal
  - Partial deletions cause conditional phenotype
  - Truncations impairs enhancer functions, initiation, and mRNA processing.
  - Mice with $2\times \Delta_{13}$ CTD: high neonatal lethality + born smaller

- **Not essential in vitro**
  - Not required for GTF-mediated initiation and RNA synthesis in vitro.
  - CTD not part of the catalytic essence of RNAPII; it must perform other functions.

- **Different promoters - different dependence on CTD**
  - Yeast CTD-deletion $n=27\rightarrow 11$, effect: $\text{GAL4} \downarrow$ $\text{HIS4} = \downarrow$
The largest subunit of yeast RNAP II

\[
\text{Tyr}_1-\text{Ser}_2-\text{Pro}_3-\text{Thr}_4-\text{Ser}_5-\text{Pro}_6-\text{Ser}_7
\]
CTD is highly phosphorylated

- Full of residues that can be phosphorylated
  - Tyr₁-Ser₂-Pro₃-Thr₄-Ser₅-Pro₆-Ser₇
- Reversible phosphorylation occurs on both Ser and Tyr
- Creates different forms of RNAPII
  - RNAPII₀ - hyperphosphorylated (Mr=240k)
    - ≈ 50 phosphates (one per repeat)
    - Abl- phosphorylated in vitro ≈30 fosfat
  - RNAPIIIA - without phosphate (Mr=214k)
  - RNAPIIIB - with CTD deleted
CTDs phosphorylation changes during the transcription cycle

- **Function of RNAPIIA ≠ RNAPIIO**
  - Assembly of pre-initiation complexes (PIC): only non-phosphorylated RNAPIIA
  - Elongation complex: only hyperphosphorylated RNAPIIO

- **Phosphorylation status changes during the transcription cycles**
  - Phosphorylation occurs after PIC assembly
  - Dephosphorylation - on free polymerase or upon termination
CTD-phosphorylation changes during the transcription cycle

PIC assembly

RNAP II

RNAP IIA

RNAP IIO

CTDK

initiering

elongering

defosforylering

CTDP

fri RNAP IIO

fosforylering

fri RNAP IIO
CTD - properties = phosphorylation + protein binding

A major function of the CTD is to serve as a binding scaffold for a variety of nuclear factors, which factors bind is determined by the phosphorylation patterns on the CTD repeats.
CTD is binding several proteins

- **Mediator:** CTD binds SRBs - suppressors of RNA pol. B
  - genetic evidence
  - mutated SRB proteins may abolish the effect of CTD deletions
  - SRBs = components of the Mediator - more later

- **GTFs**
  - TBP
  - TFIIF (74 kDa subunit)
  - TFIIE (34 kDa subunit)

- **Several proteins involved in pre m-RNA processing**
  - Many CTD-binding proteins have been identified having important functions in splicing and termination - more later
CTD structure?

CTD peptide structure
- probably flexible able to adopt several conformations
- shown as a coil, with alternating $\beta$-turns (cyan) and extended regions (pink).

Extended conformation
- A fully phosphorylated CTD is likely to extend multiple diameters out from the globular portion of RNAPII (a stretched-out yeast CTD would extend $\sim 650 \, \text{Å}$, and the mammalian CTD is twice as long; the diameter of the globular portion of the RNAPII enzyme is $\sim 150 \, \text{Å}$

CTDs function

1. Function: in initiation - recruitment
   - Role in recruitment of RNAPII to promoters
     - Only RNAPIIA can initiate PIC-assembly
     - Interactions with GTFs (more next lecture)

2. Function: in promoter clearance
   - Def: The process whereby RNAPII undergoes the transition to hyperphosphorylated elongation modus
   - CTD phosphorylation disrupts interactions and RNAPII gets free from PIC

3. Function: in elongation and pre-mRNA processing
   - CTD phosphorylation creates novel interactions with elongation and processing factors playing a role in pre-mRNA maturation
Regulation by CTD kinases/phosphatases - the logic

- **CTD kinases**
  - specific for free RNAPII ➔ repression
  - specific for assembled RNAPII ➔ activation

- **CTD phosphatases**
  - specific for free RNAPII ➔ activation
  - specific for template associated RNAPII ➔ repression
Regulated CTD-phosphorylation

- **Inhibitory**
  - CTDP
  - PIC assembly
  - RNAPII
  - klart til ny assembly
  - defosforligering

- **Stimulating**
  - CTDP
  - RNAPII
  - RNAPIIO
  - initiering
  - elongering

- **Stimulating**
  - CTDP
  - RNAPII

- **Inhibition (pausing)**
  - RNAPIIO
  - CTDP

Regulering via CTD kinaser og fosfataser

Inhibering

Stimulering
CTD kinases

- Several CTD-kinases = Cdk´s
  - Four of the putative CTD kinases are members of the cyclin-dependent kinase (CDK)/cyclin family whose members consist of a catalytic subunit bound to a regulatory cyclin subunit.
Candidate CTD-kinases

- **CTD kinase in TFIIH - positive action (more in next lecture)**
  - Good candidate with respect to timing and location
  - A multisubunit factor recruited in the last step of PIC assembly
  - TFIIH associated CTD-Kinse = MO15/CDK7 (vertebrates) = KIN28 (yeast)
  - Phosphorylates Ser5 in CTD

- **CTD kinase Srb10/11 - negative action**
  - cyclin-cdk pair (SRB10/11)
  - Conserved - human SRB10/11 also called CDK8-cyclin C
  - Isolated as a ΔCTD supressor - but recessive and with negative function in trx
  - Phosphorylates Ser5 in CTD
  - Unique by phosphorylating CTD of free RNAPII - hence negative effect on trx

- **Other candidates**
  - in vitro - CTD is substrate for several kinases
  - CDK9: component of P-TEFb, a positive-acting elongation factor
  - MAP kinases (ERK type),
  - c-Abl Tyr-kinase,
Pattern of serines phosphorylated changes during the transcription cycle

- The phosphorylation pattern changes during transcription
  - Ser 5 phosphorylation is detected mainly at promoter regions (initiation)
  - Ser 2 phosphorylation is seen only in coding regions (elongation)

\[
\text{RNAP} \quad (Y_1S_2P_3T_4S_5P_6S_7) \quad \text{P} \quad \text{Initiation}
\]

\[
\text{RNAP} \quad (Y_1S_2P_3T_4S_5P_6S_7) \quad \text{P} \quad \text{Elongation}
\]
Pattern of serines phosphorylated changes during the transcription cycle

- The phosphorylation pattern changes during transcription
  - Ser 5 phosphorylation is detected mainly at promoter regions (initiation)
  - Ser 2 phosphorylation is seen only in coding regions (elongation)
Four states of PCTD

CTD exists in at least four major phosphorylation states.

1. non-phosphorylated
   - RNAPII at a promoter initially carries a largely unphosphorylated CTD, and the enzyme is associated with a set of factors, such as Mediator.

2. Phospho-Ser5 state
   - Early in the transition from preinitiation to elongation, the CTD is phosphorylated on Ser5 residues; 5-end processing factors now bind.

3. Double phospho (Ser2, Ser5) state
   - After initiation, an elongation-phase kinase (CTDK-I in yeast; P-TEFb in metazoa) modifies mainly Ser2 residues to generate elongation-proficient RNAPII; elongation-related factors such as Set2 bind to the CTD in this third state of phosphorylation.

4. Phospho-Ser2 state
   - Near the 3 end of the gene CTD phosphorylation is dominated by Ser2P residues; consistent with the localization of 3-end processing factors.
Phosphorylated CTD (PCTD) not homogeneous

- Probably mixed forms

Example CTDs

Repeats: \{ NP, 5P, 2,5P, 2P \}
CTD phosphatases

- **First CTD phosphatase characterized = FCP1**
  - Fcp1p is necessary for CTD dephosphorylation in vivo
  - Yeast cells with temperature-sensitive mutations have severe defects in poly(A)+ mRNA synthesis at the nonpermissive temperature
- **FCP1 dephosphorylates Ser2 in CTD**
- **Function - elongation and recycling**
  - Human FCP1 can stimulate elongation by RNAPII
  - FCP1 presumably helps to recycle RNAP II at the end of the transcription cycle by converting RNAP IIO into IIA for another round of transcription.
- **Other CTD phosphatases specific for Ser5**
  - SCPs - a family of small CTD phosphatases that preferentially catalyze the dephosphorylation of Ser5 within CTD. SCP1 may play a role in transition from initiation/capping to processive transcript elongation.
  - Ssu72, a component of the yeast cleavage/polyadenylation factor (CPF) complex, is a CTD phosphatase with specificity for Ser5-P. Ssu72 may have a dual role in transcription: in recycling of RNAP II and in trx termination.
CTD phosphatase

- FCP1 is phosphorylated - regulatory target?
  - FCP1 is phosphorylated at multiple sites in vivo. Phosphorylated FCP1 is more active in stimulating transcription elongation than the dephosphorylated form.
  - Ex: The peptidyl-prolyl isomerase Pin1 influences the phosphorylation status of the CTD by inhibiting the CTD phosphatase FCP1 and stimulating CTD phosphorylation by cdc2/cyclin B.

- FCP1 is disease related
CTD kinase and phosphatase specificities

RNAP

P-TEFb (CDK9)  TFIIH (CDK7/ Kin28)

Srb10 (CDK8)

(Y₁S₂P₃T₄S₅P₆S₇)

Fcp1  SCPs  Ssu72
A hypothetical RNAPII elongation megacomplex

- CTD coordinating functions associated with transcription

Examples of questions for the exam

- **Structure function of RNAPII**
  - RNA polymerase II (RNAPII) is the key enzyme in the process of transcription. Describe briefly its overall structural design and mention some key regions in the enzyme including the three main channels and their function.

- **CTD - function, binding scaffold, and phosphorylation target**
  - The largest subunit of RNAPII contains a particular repeat-structure. Describe briefly its composition, modification and function including how it changes during the transcription cycle.